UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria dell'Informazione
Corso di Laurea magistrale in Ingegneria delle telecomunicazioni

Tesi di Laurea

Optimal multi-user transmission strategies for Energy Harvesting devices

Strategie ottime di trasmissione multi-utente per dispositivi Energy Harvesting

Laureando:
Davide Del Testa

Relatore:
Prof. Michele Zorzi

Anno Accademico 2012-2013
8 Ottobre 2013

Ai miei genitori

Abstract

In this thesis a pair of energy harvesting devices (EHDs) is considered, whose state at any given time is determined by the energy level and an importance value, associated to the transmission of a data packet to the receiver at that particular time. For each device, the objective is to optimize the transmission strategy of the two nodes over a shared wireless channel, with the goal of maximizing the long-term average importance of the transmitted data.

In the first part of this work, under the assumption of i.i.d. Bernoulli energy arrivals and a collision channel model, a central controller with perfect information on the energy level and packet importance of both nodes is examined, showing the optimality of threshold policies with respect to the data importance level, *i.e.*, the sensor nodes should report only data with an importance value above a given threshold. In addition, numerical results are provided, in order to evaluate the impact on the performance of factors such as the battery capacity size and the energy harvesting rate, as well as the interactions between the two nodes.

In the final part, recognizing that the estimate of the energy level of the batteries employed in real-world EHDs is not trivial, the focus is shifted on the design of optimal operation policies maximizing the long-term reward under imperfect knowledge of the State-Of-Charge (SOC) of the two devices. In particular, the case of a four-state controller only knowing if each SOC is HIGH or LOW is studied, providing numerical results to be compared with the case of a controller with perfect SOC knowledge.

Sommario

Questa tesi considera una coppia di dispositivi Energy Harvesting, il cui stato, in ogni istante temporale, è determinato dal livello di energia e da un valore di importanza associato alla trasmissione di un pacchetto di dati verso il ricevitore. L'obiettivo di ogni dispositivo è quello di ottimizzare la strategia di trasmissione dei due nodi attraverso un canale wireless condiviso, al fine di massimizzare l'importanza asintotica media dei dati trasmessi.

Nella prima parte della tesi, assumendo che il recupero dell'energia sia modellato da un processo stocastico bernoulliano i.i.d. e che trasmissioni simultanee collidano, si mostra l'ottimalità di una politica a soglia rispetto al livello di importanza, utilizzando un controllore centrale con conoscenza perfetta del livello di energia e dell'importanza del pacchetto di ogni sensore. Questo significa che i nodi debbano trasmettere soltanto i dati il cui valore di importanza si trova al di sopra di una data soglia. Vengono inoltre presentati risultati numerici con cui valutare l'interazione tra i due dispositivi e l'impatto sulle performance di fattori come la capacità della batteria e il tasso di recupero dell'energia.

Nella parte finale, dopo aver riconosciuto le difficoltà pratiche che impediscono una corretta stima del livello energetico delle batterie utilizzate nei sensori reali, vengono individuate delle politiche di trasmissione ottime che portino ad una massimizzazione del guadagno nel caso di conoscenza imperfetta dello stato energetico dei due dispositivi. In particolare, viene studiato un controllore a quattro stati, in grado di sapere soltanto se lo stato di carica di ogni nodo è alto o basso. Anche in questo caso vengono forniti dei risultati numerici che possano permettere un confronto con un controllore dotato di conoscenza perfetta della carica.

Contents

List of Figures

Chapter 1

Introduction

Recently, multifunctional sensor nodes capable of communicating over short distances have been developed, thanks to the advances in micro-electro-mechanical systems technology, digital electronics and wireless communications. The possibility of producing such sensors in a low-power, low-cost and small-size way, leverages the idea of sensor networks based on collaborative efforts of a large number of nodes.
These small devices, also named motes, are able to sense physical quantities (position, temperature, humidity, etc), to process data and to communicate to each other (see Figure 1.1). The sensor nodes constituting the network can be deployed very close to the phenomenon and their position does not need to be predetermined, thus allowing random deployment in inaccessible or disaster areas [2].

Sensor networks can be used in a lot of applications whose realizations require

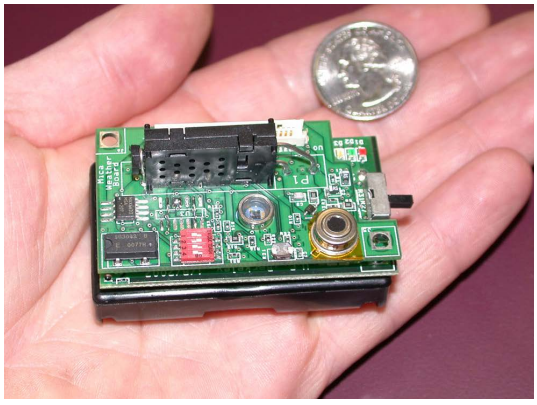

Figure 1.1: A practical sensor node [1]

wireless ad hoc networking techniques. However, most of the algorithms used

in ad hoc networks are not well suited for the requirements of the sensors. In fact, the most important differences between the two types of networks are [3]:

- sensor nodes are typically densely deployed;
- the number of nodes constituting a sensor network can be much bigger than that of an ad hoc network;
- sensor nodes are prone to failures;
- sensor nodes mostly use broadcast communications whereas ad hoc networks are based on point-to-point communications;
- the topology of a sensor network can frequently change due to mobility or damages;
- nodes are limited with respect to energy, memory and computational power.

Due to these reasons, this type of network needs algorithms realized specifically to manage data communication and routing.

The communication is usually performed in an asymmetric way: nodes send the data to one or more special nodes, called *sink* (or base station), whose aim is to collect data. The base station is a component of the WSN with much more computational power, memory, energy and communication resources, which acts as a gateway between sensor nodes and the end user, typically forwarding data from the WSN on to a server. A transmission can be initialized autonomously by the sensor (if a certain event occurs) or by the sink (by sending a *query* towards a specified node). On the other hand, it is possible to exploit the great density of nodes performing multihop communication, which is able to consume less power than the traditional single hop strategy (Figure 1.2). In addition, this technique can lower the transmission power levels (highly desirable in most scenarios) and overcome some of the signal propagation effects experienced in long-distance wireless communications.

One of the most important constraints on sensor nodes is the low-energy consumption requirement: a node generally carries a limited and irreplaceable power source. As a result, while traditional wired networks aim to meet high quality of service constraints, sensor network protocols have to focus primarily on power conservation: a permanent trade-off between network lifetime and transmission throughput has to be taken into account.

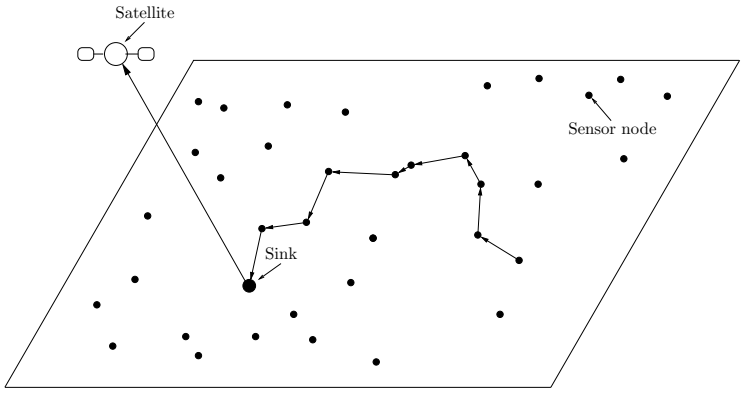

Figure 1.2: A sensor network with multihop transmission

1.1 Applications

A sensor network can be seen as a set of different types of sensors able to sense a wide variety of ambient conditions like pressure, humidity, temperature, lighting, noise but also vehicular movement, soil makeup, mechanical stress on attached objects and also the presence or absence of certain kinds of objects [4]. Moreover, nodes can be used for event detection, continuous sensing or local control of actuators useful in many application areas.

1. **Environmental**
   The term environmental sensor network has evolved to cover many applications of WSNs to earth science research [5]. Some of the major fields are listed below.

   - Air quality monitoring: in urban areas the degree of pollution of the air needs to be frequently measured in order to protect people from any kind of damage due to air pollution.
   - Natural disaster prevention: wireless sensor networks can prevent the consequences of natural disasters like earthquakes or floods. For instance, wireless nodes have successfully been installed in rivers to monitor the water levels in real time.
   - Forest fire detection: a network of sensor nodes can be deployed in a forest to detect when a fire has started. Nodes can be capable of measuring temperature, gases and humidity produced by fires among the vegetation. Thanks to WSNs, fire fighters would be able to early detect a fire and track its spreading.

- Water quality monitoring: water properties in rivers, lakes, dams and oceans, as well as underground reserves [6], can be monitored avoiding manual data retrieval in difficult-access locations.

2. **Military**
   The possibility to easily and rapidly spreading as well as self-organization and tolerance to damages make a sensor network a promising technique towards military application. Since sensor networks are based on dense deployment of low cost and disposable nodes, destruction of some nodes by enemy actions does not damage military operations as much as the destruction of a traditional sensor. A possible application is the monitoring of friendly forces, equipment and ammunition: every troop, vehicle and critical device could be attached with a small sensor to report its status. These reports could be gathered in sink nodes and sent to the troop leader or directly forwarded to the upper level of the hierarchy together with the data from the other units. Different applications are battlefield surveillance, battle damage assessment as well as nuclear, biological and chemical attack detection.

3. **Health**
   Some applications in this field are addressed to provide an interface for disabled people, to monitor human physiological data, to administer drugs in hospitals (avoiding the chance to prescribe the wrong medications to patients) and also to track doctors and patients inside hospitals [2, 7].

4. **Home**
   Sensor nodes inside the domestic appliances can interact with each other and with the external network by an Internet point. This could allow users to easily manage home appliances locally or remotely.

5. **Others**
   Some other commercial applications are virtual keyboards, interactive toys and museums, factory process control and automation, robot control and guidance in automatic manufacturing environments. Finally, also local control of actuators and vehicle tracking and detection are today available.

1.2 Characteristics of a sensor network

1.2.1 Scalability

A sensor network can be composed by thousands of nodes, and, depending on the application, this number may reach an extreme value of millions. Consequently, the implemented algorithms must be able to cope with such a heterogeneous framework. The number of nodes in a region is connected to the node density of the network, that in [7] is calculated as:

$$\rho(r) = \frac{N\pi r^2}{A},$$

where N is the number of nodes in region A and r is the radio transmission radius. Practical densities that have been used, depending on the particular application [8–10], are shown in Table 1.1.

Application	Number of nodes	Area
Machine diagnosis	300	5×5 $[m^2]$
Habitat monitoring	$25 - 100$	grid of observation
Person	~ 100	glasses, shoes, clothing
Vehicle tracking	10	5×5 $[m^2]$
Home appliances	24	home

Table 1.1: Sensor nodes densities per application

1.2.2 Fault tolerance

Fault tolerance is the ability to maintain sensor network functionality without any interruption in case of node failures. Possible arising events can be lack of power, environmental interference or physical damages. In [11], the occurrence of faults $R_k(t)$ is modelled using a Poisson process: the probability of not having a failure at node k in the time interval $(0, t)$ is

$$R_k(t) = e^{-\lambda_k t},$$

where λ_k and t are the failure rate of sensor node k and the time period, respectively.
Clearly, algorithms and protocols can be designed based on the level of fault tolerance required by the application: if the environment has no interference,

the protocol can be more relaxed. However, if the working area is the battlefield, fault tolerance must be higher, as sensed data are critical and nodes can be destroyed by hostile actions.

1.2.3 Production costs

An important constraint on a sensor node is its cost: since sensor networks consist of a large number of nodes, if the cost of this type of network is greater than that of a traditional one, then the sensor network is not cost-justified. Consequently, the cost of each sensor node has to be maintained low: in [12], it is claimed that it should be less than 1$ for the network to be feasible. However, this is a very challenging issue, since just a Bluetooth radio, which is only one of the components listed in Section 1.4, currently costs 10$ [13].

1.2.4 Topology

As sensor networks consist of up to thousand of nodes, topology issues have to be considered. Nodes can be either thrown together (for example by a plane or a rocket) or individually placed by a human or a robot. In any case, the initial deployment has to reduce the installation cost, to eliminate the need for any pre-organization and pre-planning and to facilitate self-connection and fault tolerance. Commonly, it is assumed that sensors are distributed according to a Poisson process, with fixed intensity λ. If the transmission radius is r, it is possible to transfer a message between two sensors whose distance is less than r. Figure 1.3 shows a two-dimensions network, where the dots are sensor nodes and the circles represent half their transmission radius. If two circles overlap, the communication can be efficiently established and the nodes are connected. Thanks to multihop transmission it is possible to transfer a message from A to B if there is a path of connected circles between A and B.

1.2.5 Power consumption

Some of the most important tasks of a sensor node are the event detection, the quick processing of local data and their transmission. Thus, power consumption can be divided into three areas: sensing, communication and data processing.

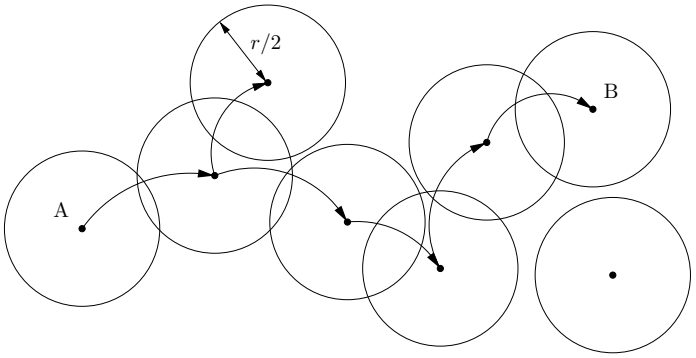

Figure 1.3: Multihop transmission between A and B

Sensing

Sensing power normally varies with the application: clearly, infrequent sensing generally consumes less power than constant event monitoring. In addition, it must be noted that sensing power is also highly influenced by the complexity of event detection: higher environmental noise levels can cause corruption and increase detection complexity.

Communication

In a sensor node the maximum quantity of energy is consumed for data communication, which involves both data transmission and reception. In fact, it has been proved that transmission cost is nearly the same as in reception in the case of short range communication. It is worth noting that these figures include also the circuitry start-up power consumption: as the transmitted packet size is reduced, this element starts to dominate the active power consumption. As a consequence, it may not be efficient to turn the transceiver ON and OFF too often.
[8] presents a formulation for the radio power consumption P_c as:

$$P_c = N_T \left[P_T \left(T_{ON} + T_{st} \right) + P_{out} \left(T_{ON} \right) \right] + N_R \left[P_R \left(R_{ON} + R_{st} \right) \right],$$

where P_T and P_R are the powers consumed by the transmitter and the receiver, P_{out} is the output power of the transmitter, T_{ON} and R_{ON} are the transmitter and receiver ON time, T_{st} and R_{st} are the transmitter and receiver start-up times and N_T and N_R are the numbers of times transmitter and receiver are switched on per unit time. Note that $T_{ON} = L/R$, L being the packet size and

R the data rate.

Data processing

With respect to data communication, the energy required for data processing is much lower. In [14], it is demonstrated that, in a generic WSN environment, the energy cost of transmitting 1 KB at a distance of 100 m is approximately the same as that of executing 3 million instructions by a 100 million instructions per second processor. Consequently, instead of sending the raw data, sensor nodes can use their processing abilities to locally carry out simple computations and transmit only the required and partially processed data.

Being a micro-electronic device, a wireless sensor node can only be provided with a limited power source. Due to the fact that, in some particular inaccessible scenarios, battery replenishment is impossible, sensor node lifetime is highly dependent on battery lifetime. As each node can play both the roles of data source and data router, for example in a multihop transmission strategy, the malfunctioning of just a few nodes can cause huge topological changes and might imply re-routing of packets and the re-organization of the network.
In traditional ad hoc networks, where power sources are assumed to be replaceable by the users, power consumption is considered a prominent design factor, but not of primary importance: the emphasis is devoted to QoS provisioning rather than power efficiency. On the other hand, in sensor networks, the latter is a very important performance metric, directly influencing the network lifetime: this thesis deals with the design of specific protocols that are designed to appropriately trade off power consumption and transmission throughput.

1.3 Harvesting capabilities

Energy harvesting (EH), also known as energy scavenging, is the process by which ambient energy is captured and, if necessary, stored to provide electricity for small autonomous devices, such as satellites, mobile phones or nodes in sensor networks. EH is used for many reasons, from providing long life and no maintenance to saving costs. Examples are given in Table 1.2 [15].

Energy management is one of the main issues in WSN because it critically threats their sustainability. Since the nodes may be distributed in extensively wide and complex environments, it becomes very difficult to replace the battery: various studies have been performed to increase the lifetime of the battery

Device	Primary reason for EH
Mobile phones, e-books, laptops	Convenience: no drained batteries, never need to find a charging point or carry a charger
Wireless sensor networks	Mobile and inaccessible nodes become feasible in challenging deployments like forests or engines; support costs greatly reduced
Military equipment	Operational availability: security of use
Medical implants	Safety; intrusive procedures reduced
Health care disposables	Cost, convenience, reliability
Consumer goods and packaging	Cost, new merchandising features

Table 1.2: Reasons for Energy Harvesting

of a node by choosing the best modulation strategy [16], by exploiting power saving modes (sleep/listen) periodically [17], by reducing the number of bits to transmit [18,19], by using energy efficient routing [20,21] and MAC [22] and by using efficient transmission scheduling to take advantage of charge recovery phenomenon [23].
In addition to these, another very interesting strategy is that of exploiting Energy Harvesting techniques [24], among which solar cells are maybe the most intuitive and well-working. A list of some scavenging sources is presented below.

1.3.1 Types of scavenging

There are many free energy sources in nature [25,26]: how to harvest and store this energy efficiently in small devices is still an open topic for research.

- **Solar** The basic principle of optical collection is to absorb a large number of photons by the use of photovoltaic materials. The main disadvantage of this energy source is the great dependence on time and on solar environment exposure. Indeed during night and cloudy days sufficient incoming energy cannot be guaranteed.

- **Thermal** Thermoelectric scavenging exploiting the differences of temperature is nowadays a very well known technology. Devices of this type can be small, light and are able to work in harsh environments.

- **Motion** If nodes are subject to movements, oscillations and vibrations,

energy could be scavenged according to Faraday's law of electromagnetic induction. The main advantage of this source is that, in some particular scenarios, it could provide constant energy.

- **Electromagnetic** When a node is exposed to an electromagnetic field, energy can be drawn with the use of an inductor. Manos Tentzeris, a professor in the Georgia Tech School of Electrical and Computer Engineering, and his team state: "There is a large amount of electromagnetic energy all around us, but nobody has been able to tap into it. We are using an ultra-wideband antenna that lets us exploit a variety of signals in different frequency ranges, giving us greatly increased power-gathering capability" [27]. It is believed that the technique could provide a promising new way to power wireless sensors networks.

Thanks to EH, energy conservation does not need to be the most important concern any more, and energy efficient policies and harvesting techniques can be jointly developed. However, some issues related to scavenging are still present: harvested energy could not always be available (for example with solar cells), despite sensor nodes' constant needs, or energy generation rates could be limited and hence the energy generation profile of the harvesting source should be matched with the energy consumption profile of the sensor node, in a way avoiding energy starvation from being the main reason for the node to die (*energy neutral operation* [24]). Furthermore, it must be noted that, treating energy arrivals as a random process, the way in which harvesting occurs is random and barely predictable.

1.3.2 Paradigms for energy efficient operations

In conclusion, differently from traditional sensors, where the objective is to minimize energy consumption under a performance constraint (for example the delay [28]), with Energy Harvesting Devices, which are utilized in this thesis, the objective is the "management" of the harvested energy. Intuitively, when a finite battery is available, an EHD should judiciously perform its assigned task based on its available energy, becoming more "conservative" as its energy supply runs low to ensure uninterrupted operation, and more "aggressive" when energy is abundant, to avoid that harvested energy is wasted due to lack of storage space.

1.4 Node architecture

The overall architecture of a sensor node consists in five basic components as shown in Figure 1.4:

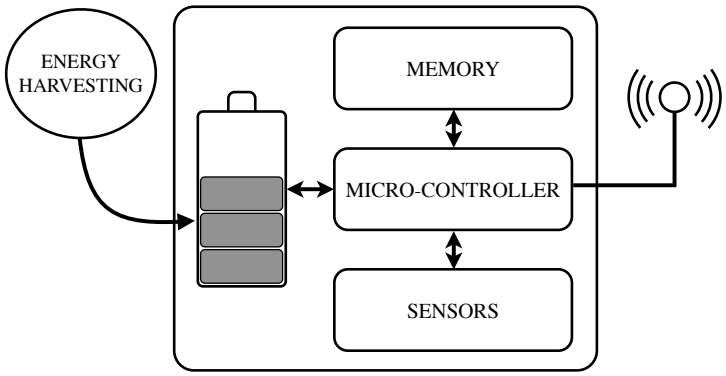

Figure 1.4: Architecture of a wireless sensor node

- The **power supply** unit of the sensor node provides power to all its components. In the majority of cases it consists of a rechargeable DC battery and can be supported by a power scavenging unit such as solar cells.
- **Micro controller** is responsible for all processing and decision making.
- **Sensing units** monitor the surrounding environment and inform the controller about what is being observed. For example, they can sense light, temperature, humidity, pressure. Sensing units are usually composed by two subunits: a sensor and an analog to digital converter (ADC). The analog signal produced by the sensor is converted to digital by the ADC, and then passed to the processing unit.
- The **Transceiver** deals with transmission and reception of the data to and from the base station. Usually RF based communication is preferred, as Infrared or Laser technologies need a direct propagation path for a correct communication.
- Sensor nodes are equipped with a programmable Flash **memory**. Usually storage capacity is limited, so the protocols that are designed for sensor networks should be simple enough to be loaded into the small available memory.

- There can also be some applications-dependent additional components, such as a **location finding system** (GPS).

1.5 Aim of the thesis

This thesis focuses on the case of a Wireless Sensor Network (WSN) consisting of two EHDs, which report incoming data of different "importance" levels to a receiver "RX" (also called Central Controller "CC"), with the overall goal to maximize the long-term aggregate average importance of the reported data.
A number of practical examples fall under this general model: a network of temperature sensing EHDs, where the importance is an increasing function of the temperature value, higher temperatures being the indicator of overheating or fire; an EHD which relays different priority packets in a sensor network [29]; an EHD which adjusts the packet information rate based on the channel condition to the RX, in which case the importance level corresponds to the instantaneous rate and the objective is to maximize the longterm average throughput [24].
Specifically, the model in [30] is generalized, which is devoted to the analysis and design of optimal energy management policies for the scenario with a single EHD, thus making a first step towards the analysis of more general networking scenarios for EHDs.
With respect to [31], which considers the design of decentralized multiaccess policies for EH wireless sensor networks, the majority of this work, which is derived from [32], assumes a centralized approach, where a central unit has perfect knowledge of the amount of energy currently available in the battery of each device, as well as the importance of the current data packets, and allocates transmission to either user accordingly. Such framework serves as a performance benchmark for the analysis of decentralized approaches.
Finally the last chapter, extending [33], relaxes the previous assumption of perfect knowledge of the batteries states-of-charge (SOC), and deals with the quantification of the performance degradation due to imperfect SOC knowledge, identifying the optimal amounts of energy to be drawn by the CC to maximize the chosen reward metric, in the special case of a two-equal interval energy level uncertainty.

Chapter 2

System model

A Wireless Sensor Network (WSN) consisting of two EHDs is considered, where each device is able to transmit to a unique Central Controller (CC). The time is divided into slots, and slot k corresponds to the time interval $[k, k+1), k \in \mathbb{Z}^+$. It is assumed that at every time instant k each EHD has a new data packet to send to the CC, whose duration is one slot; in addition, no data buffer is provided, so not packets that are not transmitted are discarded. The collision model used imposes that two simultaneous transmissions of the two EHDs incur a transmission failure: to prevent this from happening a centralized controller is employed, which has to allocate the transmissions of both sensors in order to allow only one of them in a single time slot.

The energy storage capability of each EHD is modeled by a buffer. As in previous works [34] and [35], the simplifying assumption that each position in the buffer can hold exactly one energy quantum and that the transmission of one data packet requires the expenditure of exactly one energy quantum is used. As a result, the maximum number of quanta that can be stored in EHD i is $e_{\max,i}$ and the set of its possible energy levels is denoted by $\mathcal{E}_i = \{0, 1, \ldots, e_{\max,i}\}$, $i = 1, 2$.
If the amount of energy quanta at time k at EHD i are denoted as $E_{i,k}$, the evolution of $E_{i,k}$ is determined by the following equation

$$E_{i,k+1} = \min\left\{[E_{i,k} - Q_{i,k}]^+ + B_{i,k}, e_{\max,i}\right\}, \tag{2.1}$$

where:

- $\{B_{i,k}\}$ is the *energy arrival process*, which models the randomness in the energy harvesting mechanism, *e.g.*, due to an erratic energy supply.

$\{B_{i,k}\}$ is assumed to be an i.i.d. Bernoulli process with mean $\bar{b}_i \in (0,1)$, independent across EHDs;

- $\{Q_{i,k}\}$ is the *action process*, whose value is one if the current data packet is transmitted by EHD i, resulting in the consumption of one energy quantum, drawn from the buffer, and zero otherwise. Clearly, due to the collision model employed and the centralized controller, $Q_{2,k} = 0$ if $Q_{1,k} = 1$, and vice versa.

Finally, it is assumed that a new energy quantum harvested in slot k can only be used at a later time instant $> k$.

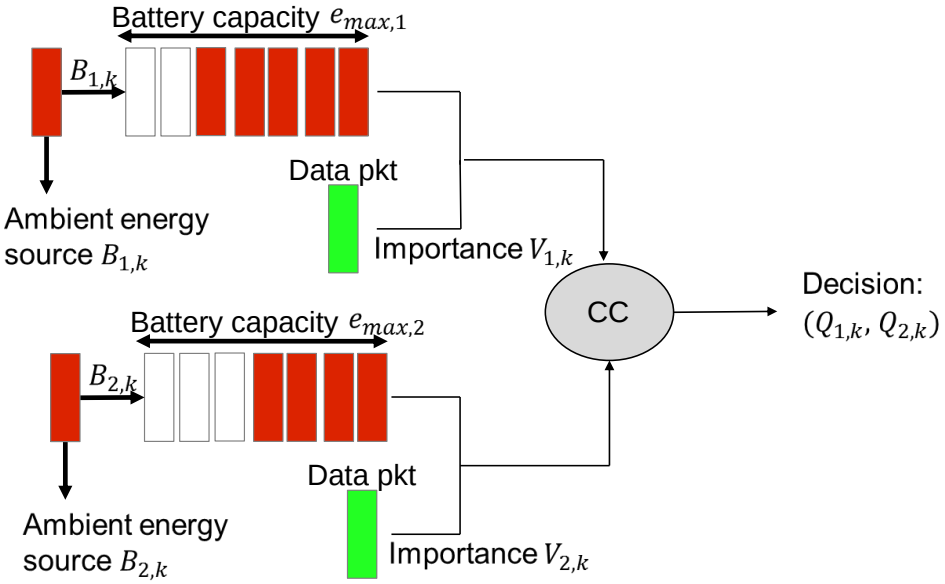

Figure 2.1: Decision process

The energy harvesting mechanism described in (2.1) and Figure 2.1 entails the following two important phenomena: *energy outage* and *energy overflow*.

Definition 1. *In slot k, for EHD i, energy outage occurs if $E_{i,k} = 0$*

Outage corresponds to the EHD running out of energy before the completion of the requested task, which happens when $Q_k > E_k$. In this case, no transmission can be performed, regardless of the importance of the current data packet, and the battery is depleted. If perfect knowledge of E_k is available at the EHD controller, then outage can always be avoided by choosing $Q_k \leq E_k$.

Definition 2. *In slot k, for EHD i, energy overflow occurs if* $(E_{i,k} = e_{\max,i}) \cap (B_{i,k} = 1) \cap (Q_{i,k} = 0)$

When energy overflow occurs, the incoming energy quantum cannot be stored in the energy buffer of EHD i, due to its finite storage capacity: as a result, since energy is lost, energy overflow potentially represents a lost transmission opportunity in the future.

At time k, the state of the system is defined as $\boldsymbol{S}_k = (E_{1,k}, E_{2,k}, V_{1,k}, V_{2,k})$, where $V_{i,k}$ is the *importance value* of the current data packet at EHD i. It is assumed that $V_{i,k} \in \mathbb{R}^+$ is a continuous random variable with probability density function (pdf) $f_{V_i}(v_i),\ v_i \geq 0$, and that $\{V_{i,k}\}$ are i.i.d. across time and EHDs (see Figure 2.2).

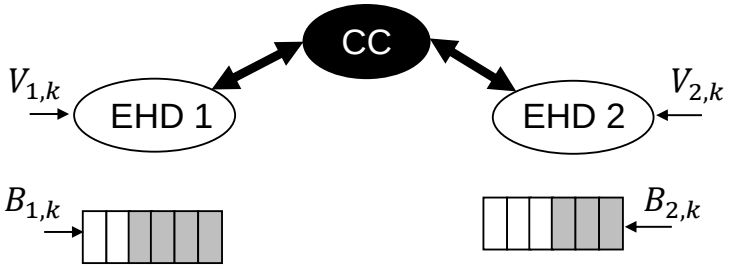

Figure 2.2: System model

2.1 Policy definition and general optimization problem

Given $\boldsymbol{S}_k$, a *policy* determines $(Q_{1,k}, Q_{2,k}) \in \{(0,0),(1,0),(0,1)\}$ at time k. Formally, a policy μ is a probability measure on the action space $\{(0,0),(1,0),(0,1)\}$, parameterized by the state $\boldsymbol{S}_k$, *i.e.*, given that $\boldsymbol{S}_k = (e_1, e_2, v_1, v_2) \in \mathcal{E}_1 \times \mathcal{E}_2 \times (\mathbb{R}^+)^2$, $\mu((i,j); \mathbf{e}, \mathbf{v})$ is the probability of drawing i and j energy quanta from EHDs 1 and 2, respectively[1]. Obviously, as collisions are avoided by a centralized controller, it results that $(Q_{1,k}, Q_{2,k}) \neq (1,1)$.

Given an initial state $\boldsymbol{S}_0$, the *long-term average reward* under policy μ is

[1]For the sake of maximizing a long-term average reward function of the state and action processes, it is sufficient to consider only state-dependent stationary policies [36].

defined as

$$G(\mu, \boldsymbol{S}_0) = \lim_{K\to\infty} \inf \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} (Q_{1,k}V_{1,k} + Q_{2,k}V_{2,k}) \,\middle|\, \boldsymbol{S}_0\right],$$

where the expectation is taken with respect to $\{B_{1,k}, B_{2,k}, Q_{1,k}, Q_{2,k}, V_{1,k}, V_{2,k}\}$ and $(Q_{1,k}, Q_{2,k})$ is drawn according to μ. The optimization problem at hand is to determine the optimal μ^* such that

$$\mu^* = \arg\max_{\mu} G(\mu, \boldsymbol{S}_0). \tag{2.2}$$

As for the single user scenario [37], it can be proved (using the Lagrangian multiplier method) that the optimal policy μ^* must have a threshold structure with respect to the importance of the current data packet. Specifically, for each pair of joint energy levels $\mathbf{e} = (e_1, e_2) \in \mathcal{E}_1 \times \mathcal{E}_2$, there exists a pair of thresholds $(v_{1,\text{th}}(e_1, e_2), v_{2,\text{th}}(e_1, e_2))$ such that

$$\begin{cases} \mu((1,0); \mathbf{e}, v_1, v_2) = 1, & v_1 > v_{1,\text{th}}(\mathbf{e}), \\ & v_1 - v_{1,\text{th}}(\mathbf{e}) \geq v_2 - v_{2,\text{th}}(\mathbf{e}), \\ \mu((0,1); \mathbf{e}, v_1, v_2) = 1, & v_2 > v_{2,\text{th}}(\mathbf{e}), \\ & v_2 - v_{2,\text{th}}(\mathbf{e}) > v_1 - v_{1,\text{th}}(\mathbf{e}), \\ \mu((0,0); \mathbf{e}, v_1, v_2) = 1, & v_1 \leq v_{1,\text{th}}(\mathbf{e}),\ v_2 \leq v_{2,\text{th}}(\mathbf{e}). \end{cases}$$

Intuitively, this means that, for the two possible transmission probability budgets $E_V[\mu((1,0); \mathbf{e}, v_1, v_2)]$ or $E_V[\mu((0,1); \mathbf{e}, v_1, v_2)]$, the optimal policy prioritizes the transmission of high over low importance data. A graphical scheme of the threshold-based transmission model can be seen in Figure 2.3. As a result, only the subset of policies with such threshold structure are henceforth considered.

For a policy with the above threshold structure the marginal transmission probability of EHD i when the joint energy level state is $\mathbf{e} = (e_1, e_2)$ is

$$\eta_i(\mathbf{e}) = \mathbb{E}[Q_{i,k} = 1 | E_{1,k} = e_1, E_{2,k} = e_2],\ i = 1, 2 \tag{2.3}$$

whereas the probability that neither node transmits is $\eta_0(\mathbf{e}) = 1 - \eta_1(\mathbf{e}) - \eta_2(\mathbf{e})$. Correspondingly, the expected reward as a function of the marginal probabilities $\eta_1(\mathbf{e})$ and $\eta_2(\mathbf{e})$ is defined as

$$g(\eta_1(\mathbf{e}), \eta_2(\mathbf{e})) = \mathbb{E}\left[Q_{1,k}V_{1,k} + Q_{2,k}V_{2,k} | E_{1,k} = e_1, E_{2,k} = e_2\right]. \tag{2.4}$$

Due to the threshold structure, the mapping between μ, $v_{\text{th},i}(\cdot)$ and $\eta_i(\cdot)$ is

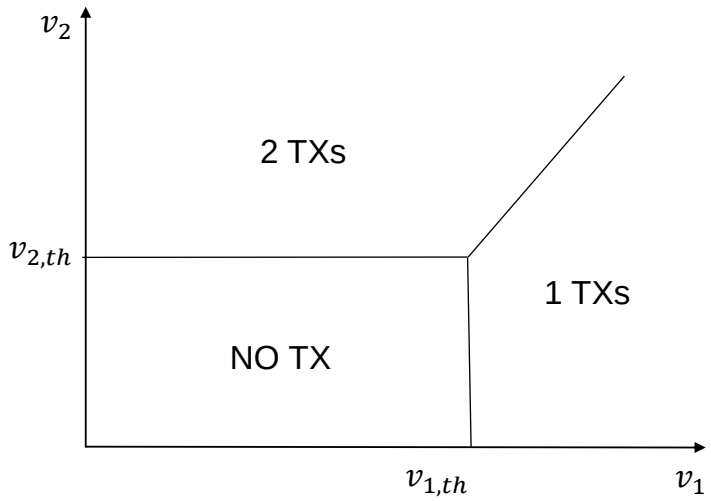

Figure 2.3: Threshold structure of the optimal policy

one-to-one, and the transition probabilities of the time-homogeneous Markov chain $\{(E_{1,k}, E_{2,k})\}$ are governed by η. As a consequence, policy μ can be referred in terms of the corresponding pair of marginal transmission probabilities (η_1, η_2). Then, assuming that (η_1, η_2) induces an irreducible Markov chain, the long-term reward does not depend on the initial state, hence the average reward becomes

$$G(\eta_1, \eta_2) = \sum_{e_1=0}^{e_{\max,1}} \sum_{e_2=0}^{e_{\max,2}} \pi_\eta(e_1, e_2) g(\eta_1(e_1, e_2), \eta_2(e_1, e_2)), \tag{2.5}$$

where $\pi_\eta(e_1, e_2)$ is the steady state distribution of the joint energy levels induced by the policy (η_1, η_2). The optimization problem in (2.2) can thus be restated as

$$(\eta_1^*, \eta_2^*) = \arg\max_{\eta_1, \eta_2} G(\eta_1, \eta_2), \tag{2.6}$$

and can be solved using standard stochastic optimization techniques, such as the Policy Iteration Algorithm (PIA) [38].

The problem of maximizing the average value of the reported data from an energy-aware replenishable sensor was formulated in [39]. However, [39] considered a continuous-time system and employed Policy Iteration to determine the optimal thresholds. [29] investigated the relaying of packets of different priorities in a network of energy-limited sensors, but did not account for energy harvesting capabilities. In [24], an EHD with a data queue was considered and sufficient stability conditions, as well as heuristic delay-minimizing policies, were derived. Other related work on EHDs includes [34, 35], which consider

variants of the system model employed in this work, but are concerned with a different performance metric, namely the probability of detection of a randomly occurring event, and [35], which proposes the use of RF-energy harvesting to enhance the performance of passive RFID systems.

Example 1. *($e_{max,1} = e_{max,2} = 1$)*
As an example, the Markov chain for the case $e_{max,1} = e_{max,2} = 1$ is drawn in Figure 2.4, where each edge shows the probability of reaching every possible energy state $\in \mathcal{E}_1 \times \mathcal{E}_2 = \{(0,0),(0,1),(1,0),(1,1)\}$. In particular, it can be seen that in the states where $e_i = e_{max,i}$, the probability of going to an energy state where still $e_i = e_{max,i}$, if no energy from EHD i is drawn, does not depend on b_i: this is a consequence of energy overflow. Furthermore, comparing this Markov chain with that for $e_{max,1} = e_{max,2} = 2$ of Figure 2.5, it is evident that the complexity of the model grows as the square of the capacity of the batteries of the two EHDs: the number of states passes from 4 to 9.

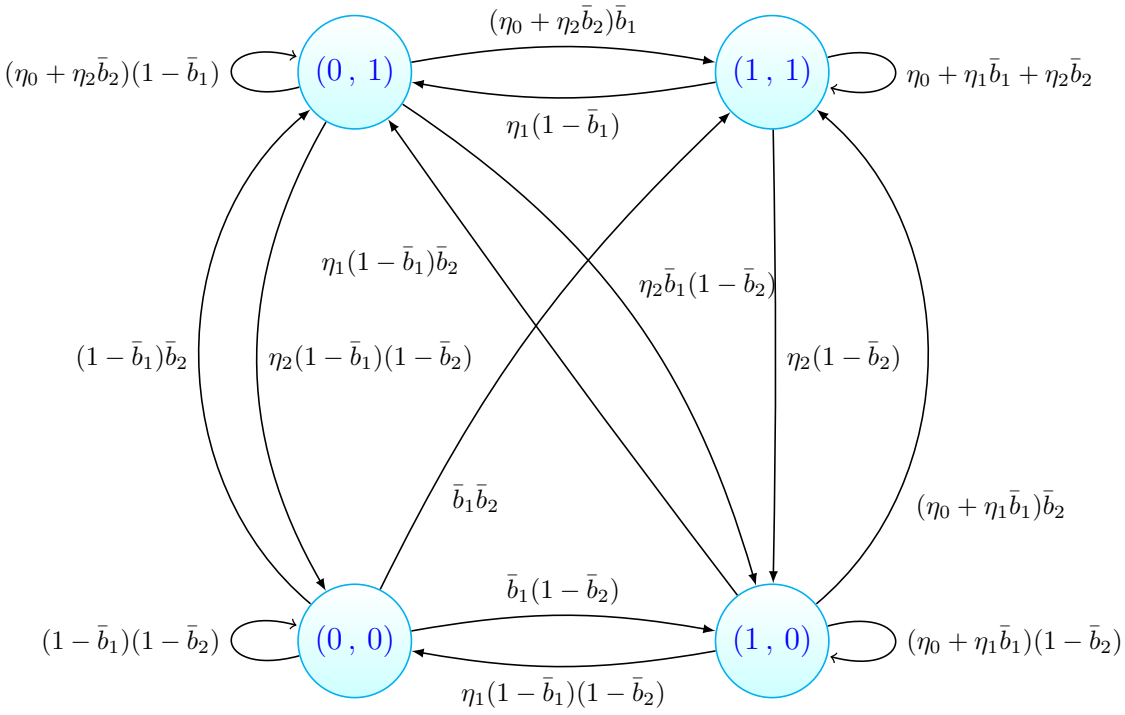

Figure 2.4: Markov chain for $e_{max,1} = e_{max,2} = 1$

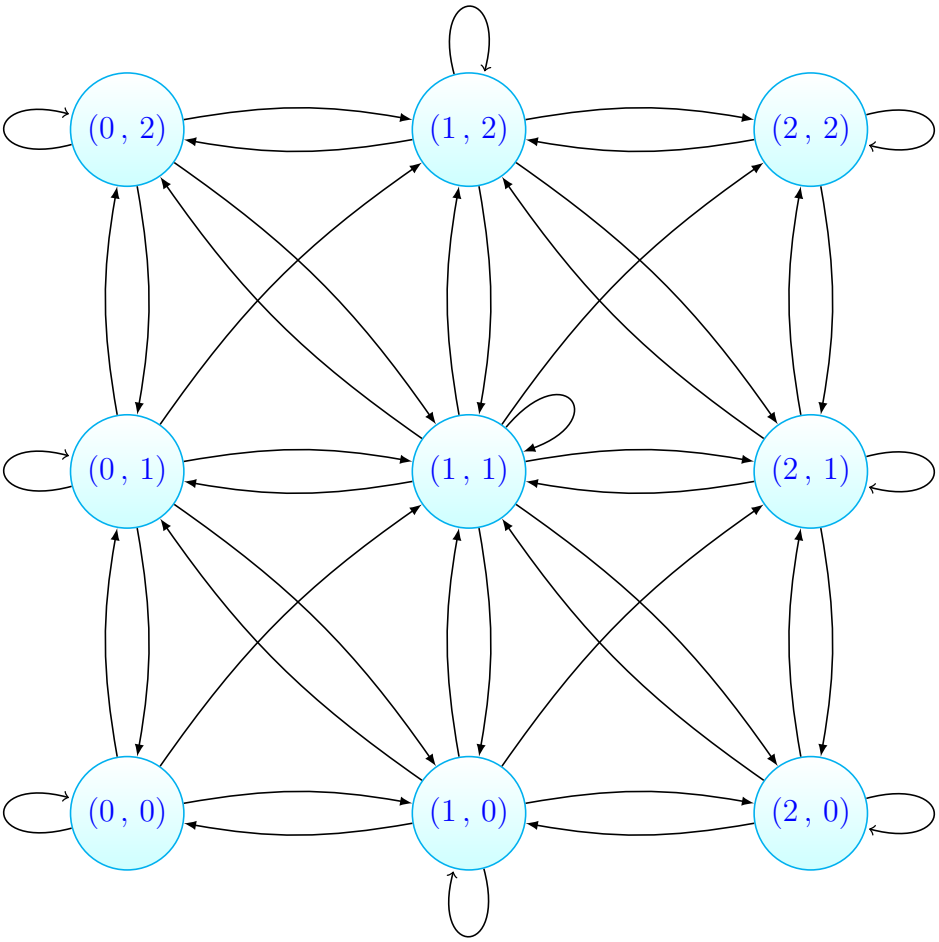

Figure 2.5: Markov chain for $e_{max,1} = e_{max,2} = 2$: labels have been omitted for clarity

2.2 Upper bound calculation for the optimal policy

It can be proved that, for any distribution of the importance values $f_{V_i}(v_i), i \in \{1,2\}$, $g(x_1,x_2)$ is a concave function of (x_1,x_2), increasing in x_1 and x_2. Thus, from (2.5),

$$G(\eta_1,\eta_2) \leq g\left(\bar{\eta}_1,\bar{\eta}_2\right), \tag{2.7}$$

where the average transmission probability of EHD i is defined as

$$\bar{\eta}_i = \sum_{e_1=0}^{e_{\max,1}} \sum_{e_2=0}^{e_{\max,2}} \pi_\eta(e_1,e_2)\eta_i(e_1,e_2). \tag{2.8}$$

Moreover, since the EH operation enforces a constraint on the maximum average transmission probability of each EHD, $\bar{\eta}_i \leq \bar{b}_i$, and the shared channel enforces the constraint $\bar{\eta}_1 + \bar{\eta}_2 \leq 1$, it results that

$$G(\eta_1,\eta_2) \leq \max_{\substack{\bar{\eta}_1 \leq \bar{b}_1 \\ \bar{\eta}_2 \leq \bar{b}_2 \\ \bar{\eta}_1+\bar{\eta}_2 \leq 1}} g\left(\bar{\eta}_1,\bar{\eta}_2\right). \tag{2.9}$$

Finally, in the special case $\bar{b}_1 + \bar{b}_2 \leq 1$ (energy-limited multiuser system), since $g\left(\bar{\eta}_1,\bar{\eta}_2\right)$ is increasing in $\bar{\eta}_i$,

$$G(\eta_1,\eta_2) \leq g\left(\bar{b}_1,\bar{b}_2\right). \tag{2.10}$$

The upper bound in (2.10) is asymptotically achievable when $e_{\max,1}, e_{\max,2} \to \infty$ by the *balanced* policy $\eta_i(e_1,e_2) = \bar{b}_i, \forall (e_1,e_2) : e_i \neq 0$, which is studied in [30] for the single EHD scenario. In fact, when $e_{\max,1}, e_{\max,2} \to \infty$, the battery of each EHD is seldom fully discharged, hence the two EHDs transmit with probability $\bar{b}_1$ and $\bar{b}_2$, respectively, in each slot. A similar consideration holds for the bound in (2.9).

Chapter 3

Maximization of the transmission rate in the low SNR regime

In this chapter, the optimization of (2.6) is considered, for the specific situation in which the long-term average transmission rate from the EHDs to the CC is to be maximized. The normalized channel gains $H_{1,k}$ and $H_{2,k}$ are assumed i.i.d. (over time and across EHDs) and exponentially distributed with unit mean, *i.e.*, with pdf $f_{H_i}(h_i) = e^{-h_i}, h_i > 0$ (Rayleigh small scale fading).
The average Signal-to-Noise-Ratio (SNR) of link i at the receiver RX is Λ_i, thus the overall SNR experienced by node i at the receiver in slot k is $\Lambda_i H_{i,k}$.
The achievable rate in slot k for EHD i is proportional to

$$V_{i,k} = \ln(1 + \Lambda_i H_{i,k}), \tag{3.1}$$

and with the above choice of the importance value, a threshold on the importance level $v_{i,\text{th}}$ corresponds to a threshold on the normalized channel gain $h_{i,\text{th}} = \frac{e^{v_{i,\text{th}}}-1}{\Lambda_i}$.
Since sensor nodes are energy constrained, a practical important case is the *low SNR regime* $\Lambda_i \ll 1$ [24]. In this case, the Shannon capacity expression (3.1) can be substituted with its linear approximation

$$V_{i,k} \simeq \Lambda_i H_{i,k}. \tag{3.2}$$

As a consequence, the maximization of the transmission rate can be performed using either (3.1) or (3.2): the latter will be treated in the next section, the former in Chapter 5.

3.1 Variables calculation

From (2.3) and (3.2), exploiting the independence between the channels, it results that:

$$\begin{aligned}\eta_0(\mathbf{e}) &= P[h_1 < h_{1,th}(\mathbf{e}), h_2 < h_{2,th}(\mathbf{e})] \\ &= \int_0^{h_{1,th}(\mathbf{e})} \int_0^{h_{2,th}(\mathbf{e})} f_{H_1,H_2}(h_1,h_2) dh_2 dh_1 \\ &= \int_0^{h_{1,th}(\mathbf{e})} \int_0^{h_{2,th}(\mathbf{e})} e^{-(h_1+h_2)} dh_2 dh_1 \\ &= \left(\int_0^{h_{1,th}(\mathbf{e})} e^{-h_1} dh_1\right)\left(\int_0^{h_{2,th}(\mathbf{e})} e^{-h_2} dh_2\right) \\ &= \left(1 - e^{-h_{1,th}(\mathbf{e})}\right)\left(1 - e^{-h_{2,th}(\mathbf{e})}\right) \end{aligned} \tag{3.3}$$

$$\begin{aligned}\eta_1(\mathbf{e}) &= P[h_1 \geq h_{1,th}(\mathbf{e}), h_1 - h_{1,th}(\mathbf{e}) \geq h_2 - h_{2,th}(\mathbf{e})] \\ &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\frac{\Lambda_1}{\Lambda_2}h_1 - \frac{\Lambda_1}{\Lambda_2}h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e})} f_{H_1,H_2}(h_1,h_2) dh_2 dh_1 \\ &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\frac{\Lambda_1}{\Lambda_2}h_1 - \frac{\Lambda_1}{\Lambda_2}h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e})} e^{-(h_1+h_2)} dh_2 dh_1 \\ &= e^{-h_{1,th}(\mathbf{e})}\left(1 - \frac{\Lambda_2}{\Lambda_1+\Lambda_2} e^{-h_{2,th}(\mathbf{e})}\right) \end{aligned} \tag{3.4}$$

and, by symmetry,

$$\eta_2(\mathbf{e}) = e^{-h_{2,th}(\mathbf{e})}\left(1 - \frac{\Lambda_1}{\Lambda_1+\Lambda_2} e^{-h_{1,th}(\mathbf{e})}\right). \tag{3.5}$$

Moreover, from (2.4), the reward is given by

$$\begin{aligned} g(\eta_1(\mathbf{e}), \eta_2(\mathbf{e})) &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\frac{\Lambda_1}{\Lambda_2}h_1 - \frac{\Lambda_1}{\Lambda_2}h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e})} \Lambda_1 h_1 e^{-(h_1+h_2)} dh_2 dh_1 + \\ &+ \int_{h_{2,th}(\mathbf{e})}^{\infty} \int_0^{\frac{\Lambda_2}{\Lambda_1}h_2 - \frac{\Lambda_2}{\Lambda_1}h_{2,th}(\mathbf{e}) + h_{1,th}(\mathbf{e})} \Lambda_2 h_2 e^{-(h_1+h_2)} dh_1 dh_2 \\ &= \Lambda_1 e^{-h_{1,th}(\mathbf{e})}(h_{1,th}(\mathbf{e}) + 1) + \Lambda_2 e^{-h_{2,th}(\mathbf{e})}(h_{2,th}(\mathbf{e}) + 1) + \\ &- \frac{\Lambda_1\Lambda_2}{\Lambda_1+\Lambda_2} e^{-(h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e}))}(h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e}) + 1). \end{aligned} \tag{3.6}$$

For the case of low SNR regime, it is possible to perform the inversion of (3.4) and (3.5), obtaining the channel thresholds $h_{1,th}(\mathbf{e}), h_{2,th}(\mathbf{e})$ corresponding to the probabilities η_0, η_1, η_2. This values are calculated in the following

Lemma.

Lemma 1. *The values of the channel thresholds $h_{1,th}(\mathbf{e}), h_{2,th}(\mathbf{e})$ corresponding to the transmission probabilities η_0, η_1, η_2 are:*

$$\begin{cases} h_{1,th}(\mathbf{e}) = \ln\left(\dfrac{(\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1) + \Delta}{2(\Lambda_1+\Lambda_2)\eta_1}\right) \\ h_{2,th}(\mathbf{e}) = \ln\left(\dfrac{2\Lambda_2}{(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) - \Delta}\right), \end{cases} \tag{3.7}$$

where

$$\Delta = \sqrt{[(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1)]^2 - 4\Lambda_2\eta_2(\Lambda_1+\Lambda_2)}.$$

Proof. The inversion of (3.3) and (3.4) leads to two possible pairs of solutions:

$$\begin{cases} h^{(i)}_{1,th}(\mathbf{e}) = \ln\left(\dfrac{(\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1) + i\Delta}{2(\Lambda_1+\Lambda_2)\eta_1}\right) \\ h^{(i)}_{2,th}(\mathbf{e}) = \ln\left(\dfrac{2\Lambda_2}{(1-\eta_0)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) - i\Delta}\right), \end{cases}$$

where $i \in \{-1, 1\}$. Of these two solutions, the pair $(h^{(-1)}_{1,th}(\mathbf{e}), h^{(-1)}_{2,th}(\mathbf{e}))$ is not feasible because either $h^{(-1)}_{1,th}(\mathbf{e}) < 0$ or $h^{(-1)}_{2,th}(\mathbf{e}) < 0$, as is shown below. In fact, the values of η_0 and η_1 leading to nonnegative $h^{(-1)}_{1,th}(\mathbf{e})$ and $h^{(-1)}_{2,th}(\mathbf{e})$ are the solutions of the system

$$\begin{cases} \ln\left(\dfrac{(\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1) - \Delta}{2(\Lambda_1+\Lambda_2)\eta_1}\right) \geq 0 \\ \ln\left(\dfrac{2\Lambda_2}{(1-\eta_0)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) + \Delta}\right) \geq 0 \end{cases}$$

i.e.

$$\begin{cases} (\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1) - \Delta \geq 2(\Lambda_1+\Lambda_2)\eta_1 \\ 2\Lambda_2 \geq (1-\eta_0)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) + \Delta \end{cases}$$

which leads to

$$\begin{cases} \Delta^2 = ((\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1-1))^2 - 4\Lambda_2\eta_2(\Lambda_1+\Lambda_2) \geq 0 \\ (\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) \geq 0 \\ (\eta_0+1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1-1) \geq 0 \\ \Delta^2 \leq ((\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1))^2 \\ \Delta^2 \leq ((\eta_0+1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1-1))^2 \end{cases}$$

whose only solution is

$$\eta_0 = 0, \quad \eta_1 = \frac{\Lambda_1}{\Lambda_1+\Lambda_2}, \quad \eta_2 = \frac{\Lambda_2}{\Lambda_1+\Lambda_2},$$

but for these values, the pair of solutions $(h_{1,th}^{(-1)}, h_{2,th}^{(-1)})$ and $(h_{1,th}^{(1)}, h_{2,th}^{(1)})$ are equivalent, as $\Delta = 0$. Otherwise, if $(\eta_0, \eta_1, \eta_2) \neq \left(0, \frac{\Lambda_1}{\Lambda_1+\Lambda_2}, \frac{\Lambda_2}{\Lambda_1+\Lambda_2}\right)$, $(h_{1,th}^{(-1)}, h_{2,th}^{(-1)})$ is not feasible.

Finally, it must be proved that (3.7) always leads to nonnegative values of $h_{1,th}(\mathbf{e})$ and $h_{2,th}(\mathbf{e})$: independently solving the inequalities

$$h_{1,th}(\mathbf{e}) = \ln\left(\frac{(\eta_0-1)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1) + \Delta}{2(\Lambda_1+\Lambda_2)\eta_1}\right) < 0$$

and

$$h_{2,th}(\mathbf{e}) = \ln\left(\frac{2\Lambda_2}{(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1) - \Delta}\right) < 0,$$

it results that $h_{1,th}(\mathbf{e}) < 0$ for $\eta_0\eta_1 < 0$ and $h_{2,th}(\mathbf{e}) < 0$ for $\eta_0 < 0$. Clearly, as $\eta_i \geq 0 \ \forall i \in \{0,1,2\}$, these conditions never occur. □

Finally, in the special case where a symmetric SNR is assumed for the two links[1], *i.e.*, $\Lambda_1 = \Lambda_2 = \Lambda$, (3.4), (3.5) and (3.6) become

$$\begin{aligned} \eta_1(\mathbf{e}) &= e^{-h_{1,th}(\mathbf{e})}\left(1 - \frac{e^{-h_{2,th}(\mathbf{e})}}{2}\right) \\ \eta_2(\mathbf{e}) &= e^{-h_{2,th}(\mathbf{e})}\left(1 - \frac{e^{-h_{1,th}(\mathbf{e})}}{2}\right) \\ g(\eta_1(\mathbf{e}), \eta_2(\mathbf{e})) &= \Lambda e^{-h_{1,th}(\mathbf{e})}(h_{1,th}(\mathbf{e})+1) + \Lambda e^{-h_{2,th}(\mathbf{e})}(h_{2,th}(\mathbf{e})+1) + \\ &\quad - \Lambda\frac{e^{-(h_{1,th}(\mathbf{e})+h_{2,th}(\mathbf{e}))}}{2}(h_{1,th}(\mathbf{e}) + h_{2,th}(\mathbf{e}) + 1), \end{aligned} \tag{3.8}$$

[1]This setting will be used in the following sections.

which shows that the SNR Λ affects the average long-term reward (2.5) as a scalar multiplication factor. Therefore, without loss of generality, in this Chapter $\Lambda = 1$ is henceforth assumed, since any other value of Λ can be obtained by scaling (notice that this is true only for low SNR, and when $\Lambda_1 = \Lambda_2$).

3.2 Algorithm implementation

The Policy Iteration Algorithm (PIA) [38] has been used to compute (2.6): starting from an initial policy $\boldsymbol{\eta}^0$, this algorithm iteratively computes the *Policy Evaluation* and *Policy Improvement* steps, until convergence. Firstly, using the initial values of $\boldsymbol{\eta}^0$, the value of (2.5) is computed, and then the *relative value function* $v_{\boldsymbol{\eta}} : \mathcal{E}_1 \times \mathcal{E}_2 \to \mathbb{R}$ is determined as the unique solution of the linear system of equations

$$\begin{cases} v_{\boldsymbol{\eta}}(0,0) = 0 \\ v_{\boldsymbol{\eta}}(\mathbf{e}) = \sum\limits_{\mathbf{j} \in \mathcal{E}_1 \times \mathcal{E}_2} P_{\boldsymbol{\eta}}(\mathbf{E}_{k+1} = \mathbf{j} | \mathbf{E}_k = \mathbf{e}) v_{\boldsymbol{\eta}}(\mathbf{j}) = g(\eta_1(\mathbf{e}), \eta_2(\mathbf{e})) - G(\eta_1, \eta_2), \end{cases} \tag{3.9}$$

$\forall \mathbf{e} \in \mathcal{E}_1 \times \mathcal{E}_2$.

Example 2. *($e_{max,1} = e_{max,2} = 1$)*
Here $\mathcal{E}_1 \times \mathcal{E}_2 = \{(0,0), (0,1), (1,0), (1,1)\}$ and for a current policy $\boldsymbol{\eta}^k = \boldsymbol{\eta}$:

$$\mathbf{v'}_{\boldsymbol{\eta}} = \mathbf{A}^{-1}\mathbf{c} \tag{3.10}$$

where

$$\mathbf{A} = \begin{bmatrix} -P_{\boldsymbol{\eta}}((0,0),(0,1)) & -P_{\boldsymbol{\eta}}((0,0),(1,0)) & -P_{\boldsymbol{\eta}}((0,0),(1,1)) \\ 1 - P_{\boldsymbol{\eta}}((0,1),(0,1)) & -P_{\boldsymbol{\eta}}((0,1),(1,0)) & -P_{\boldsymbol{\eta}}((0,1),(1,1)) \\ -P_{\boldsymbol{\eta}}((1,0),(0,1)) & 1 - P_{\boldsymbol{\eta}}((1,0),(1,0)) & -P_{\boldsymbol{\eta}}((1,0),(1,1)) \\ -P_{\boldsymbol{\eta}}((1,1),(0,1)) & -P_{\boldsymbol{\eta}}((1,1),(1,0)) & 1 - P_{\boldsymbol{\eta}}((1,1),(1,1)) \end{bmatrix}$$

$$\mathbf{v'}_{\boldsymbol{\eta}} = \begin{bmatrix} v_{\boldsymbol{\eta}}(0,1) \\ v_{\boldsymbol{\eta}}(1,0) \\ v_{\boldsymbol{\eta}}(1,1) \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} g(\eta_1(0,1), \eta_2(0,1)) \\ g(\eta_1(1,0), \eta_2(1,0)) \\ g(\eta_1(1,1), \eta_2(1,1)) \end{bmatrix} - G(\eta_1, \eta_2)$$

Secondly, in the *Policy Improvement* step, an improved policy $\boldsymbol{\eta}^{k+1}$ is de-

termined by solving, $\forall \mathbf{e} \in \mathcal{E}_1 \times \mathcal{E}_2 \setminus \{(0,0)\}$, the convex optimization problem

$$\boldsymbol{\eta}^{k+1}(\mathbf{e}) = \underset{\substack{\tilde{\eta}_1(\mathbf{e})\in[0,1]\\ \tilde{\eta}_2(\mathbf{e})\in[0,1]}}{\arg\max} \left[g(\tilde{\eta}_1(\mathbf{e}), \tilde{\eta}_2(\mathbf{e})) + \sum_{\mathbf{j}\in\mathcal{E}_1\times\mathcal{E}_2} P_{\tilde{\boldsymbol{\eta}}}(\mathbf{E}_{k+1} = \mathbf{j}|\mathbf{E}_k = \mathbf{e}) v_{\boldsymbol{\eta}^k}(\mathbf{j}) \right] \tag{3.11}$$

The process is repeated with $\boldsymbol{\eta}^{k+1}$ substituting $\boldsymbol{\eta}^k$ until $|G(\boldsymbol{\eta}^{k+1}) - G(\boldsymbol{\eta}^k)| < \epsilon_{PIA}$, with ϵ_{PIA} being a defined (small) threshold: in this case the algorithm terminates providing the policy $\boldsymbol{\eta}^k$. The following proposition establishes the validity of PIA.

Proposition 1. *Assuming that there exists an integer m such that, regardless of the policy used and the initial state, there is positive probability that the termination state will be reached after no more than m stages, the policy iteration algorithm generates an improving sequence of policies (that is, $G(\boldsymbol{\eta}^{k+1}) \geq G(\boldsymbol{\eta}^k)\ \forall k$), and terminates with an optimal policy.*

Proof. See [38]. □

Example 3. *($e_{max,1} = e_{max,2} = 1$)*
The policy improvement step for the case $e_{max,1} = e_{max,2} = 1$ consists in:

$$\begin{bmatrix} \boldsymbol{\eta}^{k+1}(0,1) \\ \boldsymbol{\eta}^{k+1}(1,0) \\ \boldsymbol{\eta}^{k+1}(1,1) \end{bmatrix} = \arg\max \left\{ \begin{bmatrix} g(\tilde{\eta}_1(0,1), \tilde{\eta}_2(0,1)) \\ g(\tilde{\eta}_1(1,0), \tilde{\eta}_2(1,0)) \\ g(\tilde{\eta}_1(1,1), \tilde{\eta}_2(1,1)) \end{bmatrix} + \mathbf{A}' \begin{bmatrix} v_{\boldsymbol{\eta}^k}(0,1) \\ v_{\boldsymbol{\eta}^k}(1,0) \\ v_{\boldsymbol{\eta}^k}(1,1) \end{bmatrix} \right\}$$

where

$$\mathbf{A}' = \begin{bmatrix} P_{\tilde{\boldsymbol{\eta}}}((0,1),(0,1)) & P_{\tilde{\boldsymbol{\eta}}}((0,1),(1,0)) & P_{\tilde{\boldsymbol{\eta}}}((0,1),(1,1)) \\ P_{\tilde{\boldsymbol{\eta}}}((1,0),(0,1)) & P_{\tilde{\boldsymbol{\eta}}}((1,0),(1,0)) & P_{\tilde{\boldsymbol{\eta}}}((1,0),(1,1)) \\ P_{\tilde{\boldsymbol{\eta}}}((1,1),(0,1)) & P_{\tilde{\boldsymbol{\eta}}}((1,1),(1,0)) & P_{\tilde{\boldsymbol{\eta}}}((1,1),(1,1)) \end{bmatrix}.$$

3.3 Calculation of the optimal thresholds

Using the simplified framework of (3.8), it is now possible to analytically obtain the optimal values $\hat{h}_{1,th}(\mathbf{e})$ and $\hat{h}_{2,th}(\mathbf{e})$ for every energy state $(e_1, e_2) \in \mathcal{E}_1 \times \mathcal{E}_2$, calculating the partial derivatives of the inner argument of (3.11)[2]:

$$J = g(\tilde{\eta}_1(\mathbf{e}), \tilde{\eta}_2(\mathbf{e})) + \sum_{\mathbf{j}\in\mathcal{E}_1\times\mathcal{E}_2} P_{\tilde{\boldsymbol{\eta}}}(\mathbf{E}_{k+1} = \mathbf{j}|\mathbf{E}_k = \mathbf{e}) v_{\boldsymbol{\eta}^k}(\mathbf{j}).$$

[2]In this section the dependence on $(\mathbf{e})$ is dropped for notational convenience.

States $(0, e_2),\ e_2 \neq 0, e_{max,2}$

$$
\begin{aligned}
J &= e^{-h_{2,th}} \left(1 - \frac{e^{-h_{1,th}}}{2}\right) A + \left(1 - e^{-h_{1,th}}\right)\left(1 - e^{-h_{2,th}}\right) B + \\
&\quad + e^{-h_{1,th}}(h_{1,th} + 1) + e^{-h_{2,th}}(h_{2,th} + 1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th} + h_{2,th} + 1) \\
\frac{\partial J}{\partial h_{2,th}} &= e^{-h_{2,th}} \left[\left(\frac{e^{-h_{1,th}}}{2} - 1\right) A + (1 - e^{-h_{1,th}})B + \frac{e^{-h_{1,th}}(h_{1,th} + h_{2,th})}{2} - h_{2,th}\right] \\
\hat{h}_{1,th} &= +\infty \\
\hat{h}_{2,th} &= \frac{e^{-h_{1,th}} h_{1,th} + (e^{-h_{1,th}} - 2)A + 2(1 - e^{-h_{1,th}})B}{2 - e^{-h_{1,th}}}
\end{aligned}
$$

with

$$
\begin{aligned}
A &= (1 - \bar{b}_1)(1 - \bar{b}_2)v(0, e_2 - 1) + \bar{b}_1(1 - \bar{b}_2)v(1, e_2 - 1) + 1 - \bar{b}_1)\bar{b}_2 v(0, e_2) + \\
&\quad + \bar{b}_1\bar{b}_2 v(1, e_2) \\
B &= (1 - \bar{b}_1)(1 - \bar{b}_2)v(0, e_2) + \bar{b}_1(1 - \bar{b}_2)v(1, e_2) + (1 - \bar{b}_1)\bar{b}_2 v(0, e_2 + 1) + \\
&\quad + \bar{b}_1\bar{b}_2 v(1, e_2 + 1)
\end{aligned}
$$

States $(e_1, 0),\ e_1 \neq 0, e_{max,1}$

$$
\begin{aligned}
J &= e^{-h_{1,th}} \left(1 - \frac{e^{-h_{2,th}}}{2}\right) A + \left(1 - e^{-h_{1,th}}\right)\left(1 - e^{-h_{2,th}}\right) B + \\
&\quad + e^{-h_{1,th}}(h_{1,th} + 1) + e^{-h_{2,th}}(h_{2,th} + 1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th} + h_{2,th} + 1) \\
\frac{\partial J}{\partial h_{1,th}} &= e^{-h_{1,th}} \left[\left(\frac{e^{-h_{2,th}}}{2} - 1\right) A + (1 - e^{-h_{2,th}})B + \frac{e^{-h_{2,th}}(h_{1,th} + h_{2,th})}{2} - h_{1,th}\right] \\
\hat{h}_{1,th} &= \frac{e^{-h_{2,th}} h_{2,th} + (e^{-h_{2,th}} - 2)A + 2(1 - e^{-h_{2,th}})B}{2 - e^{-h_{2,th}}} \\
\hat{h}_{2,th} &= +\infty
\end{aligned}
$$

with

$$
\begin{aligned}
A &= (1 - \bar{b}_1)(1 - \bar{b}_2)v(e_1 - 1, 0) + \bar{b}_1(1 - \bar{b}_2)v(e_1, 0) + (1 - \bar{b}_1)\bar{b}_2 v(e_1 - 1, 1) + \\
&\quad + \bar{b}_1\bar{b}_2 v(e_1, 1) \\
B &= (1 - \bar{b}_1)(1 - \bar{b}_2)v(e_1, 0) + \bar{b}_1(1 - \bar{b}_2)v(e_1 + 1, 0) + (1 - \bar{b}_1)\bar{b}_2 v(e_1, 1) + \\
&\quad + \bar{b}_1\bar{b}_2 v(e_1 + 1, 1)
\end{aligned}
$$

State $(e_{max,1}, e_{max,2})$

$$J = e^{-h_{2,th}}\left(1-\frac{e^{-h_{1,th}}}{2}\right)A + e^{-h_{1,th}}\left(1-\frac{e^{-h_{2,th}}}{2}\right)B + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)C +$$
$$+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1)$$
$$\frac{\partial J}{\partial h_{1,th}} = e^{-h_{1,th}}\left[\frac{e^{-h_{2,th}}}{2}A - \left(1-\frac{e^{-h_{2,th}}}{2}\right)B + (1-e^{-h_{2,th}})C + \frac{e^{-h_{2,th}}(h_{1,th}+h_{2,th})}{2} - h_{1,th}\right]$$
$$\frac{\partial J}{\partial h_{2,th}} = e^{-h_{2,th}}\left[\left(\frac{e^{-h_{1,th}}}{2}-1\right)A + \frac{e^{-h_{1,th}}}{2}B + (1-e^{-h_{1,th}})C + \frac{e^{-h_{1,th}}(h_{1,th}+h_{2,th})}{2} - h_{2,th}\right]$$
$$\hat{h}_{1,th} = \frac{e^{-h_{2,th}}h_{2,th} + e^{-h_{2,th}}A + (e^{-h_{2,th}}-2)B + 2(1-e^{-h_{2,th}})C}{2-e^{-h_{2,th}}}$$
$$\hat{h}_{2,th} = \frac{e^{-h_{1,th}}h_{1,th} + (e^{-h_{1,th}}-2)A + e^{-h_{1,th}}B + 2(1-e^{-h_{1,th}})C}{2-e^{-h_{1,th}}}$$

with

$$A = (1-\bar{b}_2)v(e_{max,1}, e_{max,2}-1) + \bar{b}_2 v(e_{max,1}, e_{max,2})$$
$$B = (1-\bar{b}_1)v(e_{max,1}-1, e_{max,2}) + \bar{b}_1 v(e_{max,1}, e_{max,2})$$
$$C = v(e_{max,1}, e_{max,2})$$

State $(0, e_{max,2})$

$$J = e^{-h_{2,th}}\left(1-\frac{e^{-h_{1,th}}}{2}\right)A + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)B +$$
$$+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1)$$
$$\frac{\partial J}{\partial h_{2,th}} = e^{-h_{2,th}}\left[\left(\frac{e^{-h_{1,th}}}{2}-1\right)A + (1-e^{-h_{1,th}})B + \frac{e^{-h_{1,th}}(h_{1,th}+h_{2,th})}{2} - h_{2,th}\right]$$
$$\hat{h}_{1,th} = +\infty$$
$$\hat{h}_{2,th} = \frac{e^{-h_{1,th}}h_{1,th} + (e^{-h_{1,th}}-2)A + 2(1-e^{-h_{1,th}})B}{2-e^{-h_{1,th}}}$$

with

$$A = (1-\bar{b}_1)(1-\bar{b}_2)v(0, e_{max,2}-1) + \bar{b}_1(1-\bar{b}_2)v(1, e_{max,2}-1) +$$
$$+ (1-\bar{b}_1)\bar{b}_2 v(0, e_{max,2}) + \bar{b}_1\bar{b}_2 v(1, e_{max,2})$$
$$B = (1-\bar{b}_1)v(0, e_{max,2}) + \bar{b}_1 v(1, e_{max,2})$$

State $(e_{max,1}, 0)$

$$
\begin{aligned}
J &= e^{-h_{1,th}}\left(1-\frac{e^{-h_{2,th}}}{2}\right)A + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)B + \\
&+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1) \\
\frac{\partial J}{\partial h_{1,th}} &= e^{-h_{1,th}}\left[\left(\frac{e^{-h_{2,th}}}{2}-1\right)A + (1-e^{-h_{2,th}})B + \frac{e^{-h_{2,th}}(h_{1,th}+h_{2,th})}{2} - h_{1,th}\right] \\
\hat{h}_{1,th} &= \frac{e^{-h_{2,th}}h_{2,th} + (e^{-h_{2,th}}-2)A + 2(1-e^{-h_{2,th}})B}{2-e^{-h_{2,th}}} \\
\hat{h}_{2,th} &= +\infty
\end{aligned}
$$

with

$$
\begin{aligned}
A &= (1-\bar{b}_1)(1-\bar{b}_2)v(e_{max,1}-1,0) + \bar{b}_1(1-\bar{b}_2)v(e_{max,1},0) + \\
&+ (1-\bar{b}_1)\bar{b}_2 v(e_{max,1}-1,1) + \bar{b}_1\bar{b}_2 v(e_{max,1},1) \\
B &= (1-\bar{b}_2)v(e_{max,1},0) + \bar{b}_2 v(e_{max,1},1)
\end{aligned}
$$

States $(e_{max,1}, e_2),\ e_2 \neq 0, e_{max,2}$

$$
\begin{aligned}
J &= e^{-h_{1,th}}\left(1-\frac{e^{-h_{2,th}}}{2}\right)A + e^{-h_{2,th}}\left(1-\frac{e^{-h_{1,th}}}{2}\right)B + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)C + \\
&+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1) \\
\frac{\partial J}{\partial h_{1,th}} &= e^{-h_{1,th}}\left[\left(\frac{e^{-h_{2,th}}}{2}-1\right)A + \frac{e^{-h_{2,th}}}{2}B + (1-e^{-h_{2,th}})C + \frac{e^{-h_{2,th}}(h_{1,th}+h_{2,th})}{2} - h_{1,th}\right] \\
\frac{\partial J}{\partial h_{2,th}} &= e^{-h_{2,th}}\left[\frac{e^{-h_{1,th}}}{2}A - \left(1-\frac{e^{-h_{1,th}}}{2}\right)B + (1-e^{-h_{1,th}})C + \frac{e^{-h_{1,th}}(h_{1,th}+h_{2,th})}{2} - h_{2,th}\right] \\
\hat{h}_{1,th} &= \frac{e^{-h_{2,th}}h_{2,th} + (e^{-h_{2,th}}-2)A + e^{-h_{2,th}}B + 2(1-e^{-h_{2,th}})C}{2-e^{-h_{2,th}}} \\
\hat{h}_{2,th} &= \frac{e^{-h_{1,th}}h_{1,th} + e^{-h_{1,th}}A + (e^{-h_{1,th}}-2)B + 2(1-e^{-h_{1,th}})C}{2-e^{-h_{1,th}}}
\end{aligned}
$$

with

$$
\begin{aligned}
A &= (1-\bar{b}_1)(1-\bar{b}_2)v(e_{max,1}-1,e_2) + \bar{b}_1(1-\bar{b}_2)v(e_{max,1},e_2) + \\
&+ (1-\bar{b}_1)\bar{b}_2 v(e_{max,1}-1,e_2+1) + \bar{b}_1\bar{b}_2 v(e_{max,1},e_2+1) \\
B &= (1-\bar{b}_2)v(e_{max,1},e_2-1) + \bar{b}_2 v(e_{max,1},e_2) \\
C &= (1-\bar{b}_2)v(e_{max,1},e_2) + \bar{b}_2 v(e_{max,1},e_2+1)
\end{aligned}
$$

States $(e_1, e_{max,2}),\ e_1 \neq 0, e_{max,1}$

$$J = e^{-h_{2,th}}\left(1-\frac{e^{-h_{1,th}}}{2}\right)A + e^{-h_{1,th}}\left(1-\frac{e^{-h_{2,th}}}{2}\right)B + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)C +$$
$$+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1)$$
$$\frac{\partial J}{\partial h_{1,th}} = e^{-h_{1,th}}\left[\frac{e^{-h_{2,th}}}{2}A - \left(1-\frac{e^{-h_{2,th}}}{2}\right)B + (1-e^{-h_{2,th}})C + \frac{e^{-h_{2,th}}(h_{1,th}+h_{2,th})}{2} - h_{1,th}\right]$$
$$\frac{\partial J}{\partial h_{2,th}} = e^{-h_{2,th}}\left[\left(\frac{e^{-h_{1,th}}}{2}-1\right)A + \frac{e^{-h_{1,th}}}{2}B + (1-e^{-h_{1,th}})C + \frac{e^{-h_{1,th}}(h_{1,th}+h_{2,th})}{2} - h_{2,th}\right]$$
$$\hat{h}_{1,th} = \frac{e^{-h_{2,th}}h_{2,th} + e^{-h_{2,th}}A + (e^{-h_{2,th}}-2)B + 2(1-e^{-h_{2,th}})C}{2-e^{-h_{2,th}}}$$
$$\hat{h}_{2,th} = \frac{e^{-h_{1,th}}h_{1,th} + (e^{-h_{1,th}}-2)A + e^{-h_{1,th}}B + 2(1-e^{-h_{1,th}})C}{2-e^{-h_{1,th}}}$$

with

$$A = (1-\bar{b}_1)(1-\bar{b}_2)v(e_1, e_{max,2}-1) + \bar{b}_1(1-\bar{b}_2)v(e_1+1, e_{max,2}-1) +$$
$$+ (1-\bar{b}_1)\bar{b}_2 v(e_1, e_{max,2}) + \bar{b}_1\bar{b}_2 v(e_1+1, e_{max,2})$$
$$B = (1-\bar{b}_1)v(e_1-1, e_{max,2}) + \bar{b}_1 v(e_1, e_{max,2})$$
$$C = (1-\bar{b}_1)v(e_1, e_{max,2}) + \bar{b}_1 v(e_1+1, e_{max,2})$$

States $(e_1, e_2),\ e_1 \neq 0, e_{max,1},\ e_2 \neq 0, e_{max,2}$

$$J = e^{-h_{1,th}}\left(1-\frac{e^{-h_{2,th}}}{2}\right)A + e^{-h_{2,th}}\left(1-\frac{e^{-h_{1,th}}}{2}\right)B + \left(1-e^{-h_{1,th}}\right)\left(1-e^{-h_{2,th}}\right)C +$$
$$+ e^{-h_{1,th}}(h_{1,th}+1) + e^{-h_{2,th}}(h_{2,th}+1) - \frac{e^{-(h_{1,th}+h_{2,th})}}{2}(h_{1,th}+h_{2,th}+1)$$
$$\frac{\partial J}{\partial h_{1,th}} = e^{-h_{1,th}}\left[\left(\frac{e^{-h_{2,th}}}{2}-1\right)A + \frac{e^{-h_{2,th}}}{2}B + (1-e^{-h_{2,th}})C + \frac{e^{-h_{2,th}}(h_{1,th}+h_{2,th})}{2} - h_{1,th}\right]$$
$$\frac{\partial J}{\partial h_{2,th}} = e^{-h_{2,th}}\left[\frac{e^{-h_{1,th}}}{2}A - \left(1-\frac{e^{-h_{1,th}}}{2}\right)B + (1-e^{-h_{1,th}})C + \frac{e^{-h_{1,th}}(h_{1,th}+h_{2,th})}{2} - h_{2,th}\right]$$
$$\hat{h}_{1,th} = \frac{e^{-h_{2,th}}h_{2,th} + (e^{-h_{2,th}}-2)A + e^{-h_{2,th}}B + 2(1-e^{-h_{2,th}})C}{2-e^{-h_{2,th}}}$$
$$\hat{h}_{2,th} = \frac{e^{-h_{1,th}}h_{1,th} + e^{-h_{1,th}}A + (e^{-h_{1,th}}-2)B + 2(1-e^{-h_{1,th}})C}{2-e^{-h_{1,th}}}$$

with

$$A = (1-\bar{b}_1)(1-\bar{b}_2)v(e_1-1, e_2) + \bar{b}_1(1-\bar{b}_2)v(e_1, e_2) +$$
$$+ (1-\bar{b}_1)\bar{b}_2 v(e_1-1, e_2+1) + \bar{b}_1\bar{b}_2 v(e_1, e_2+1)$$
$$B = (1-\bar{b}_1)(1-\bar{b}_2)v(e_1, e_2-1) + \bar{b}_1(1-\bar{b}_2)v(e_1+1, e_2-1) +$$
$$+ (1-\bar{b}_1)\bar{b}_2 v(e_1, e_2) + \bar{b}_1\bar{b}_2 v(e_1+1, e_2)$$
$$C = (1-\bar{b}_1)(1-\bar{b}_2)v(e_1, e_2) + \bar{b}_1(1-\bar{b}_2)v(e_1+1, e_2) +$$
$$+ (1-\bar{b}_1)\bar{b}_2 v(e_1, e_2+1) + \bar{b}_1\bar{b}_2 v(e_1+1, e_2+1)$$

Chapter 4

Numerical results in the low SNR regime

In this chapter, some numerical results for the two EHDs model are presented, in the low SNR regime $\Lambda_i \ll 1$, using the framework given by (3.3), (3.7) and (3.8), and assuming $\Lambda_1 = \Lambda_2 = 1$ as discussed in Chapter 3.

Figure 4.1, plots the long-term reward (throughput) for different values of $\bar{b}$ with $\bar{b}_1 = \bar{b}_2 = \bar{b}$. It can be seen that the reward keeps increasing in the capacity of the two batteries, until $e_{\max,1} = e_{\max,2} \simeq 20$, after which the performance saturates at a constant value. This is because, the larger the battery, the smaller the impact of energy outage and overflow, hence the better the performance. When the battery capacity becomes larger than 20, the improvement due to the decreased occurrence of overflow and outage events becomes negligible, and the performance is very close to the upper bound. An important implication of this result is that EHDs do not need to be equipped with very large energy buffers to achieve maximum performance. Moreover, as expected, the reward increases with the harvesting rate of the EHDs: however, due to the fact that the practical values of $\bar{b}_i$ are not very large, it cannot grow too much.

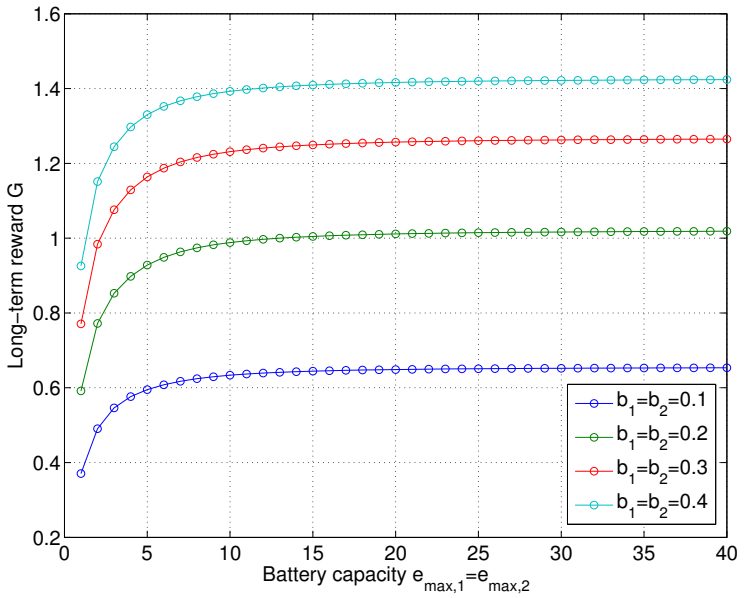

Figure 4.1: Long-term reward versus battery capacity $e_{max,1} = e_{max,2}$

Figure 4.2 shows a three-dimensional plot representing the normalized long-term reward as a function of $e_{\max,1}$ and $e_{\max,2}$, when $\bar{b}_1 = \bar{b}_2 = 0.1$. The normalization is done with respect to the upper bound $g(\bar{b}_1, \bar{b}_2)$, given in (2.10). The reward increases with the capacity of the two batteries in a "symmetric" way, in the sense that the two EHDs are interchangeable: the result achieved when $e_{\max,1} = e_1$ and $e_{\max,2} = e_2$ is the same that can be obtained when $e_{\max,1} = e_2$ and $e_{\max,2} = e_1$. This is due to the symmetry of the system, since $\bar{b}_1 = \bar{b}_2$ and $\Lambda_1 = \Lambda_2$. Furthermore, as in Figure 4.1, it can be seen that saturation occurs at $e_{\max,1} = e_{\max,2} \simeq 20$, where the upper bound (2.10) is closely approached.

The battery capacity requirements needed to achieve a specific performance are depicted in Figure 4.3, where a contour plot of Figure 4.2 has been made selecting different reward levels. 90% of the maximum performance is achieved even when the EHDs have small capacities ($e_{\max,1} = e_{\max,2} = 5$), whereas 99% can be attained only with larger batteries (in particular, in this case $e_{\max,1} = e_{\max,2} = 20$).

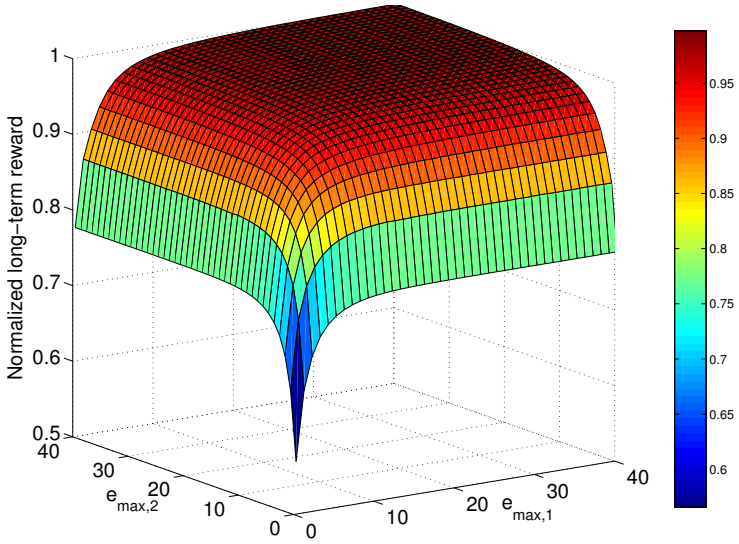

Figure 4.2: Normalized long-term reward versus battery capacities, $b_1 = b_2 = 0.1$

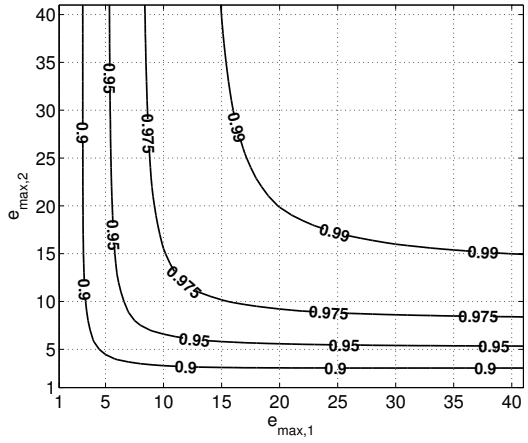

Figure 4.3: Contour plot of Figure 4.2

In Figure 4.4, the long-term reward for different values of the battery capacity $e_{\max,1}$ of the first EHD is plotted, as a function of the battery capacity $e_{\max,2}$ of the other EHD. As in Figure 4.1, the reward presents an increasing behavior until saturation occurs. Moreover, good performance is achieved when at least one user can afford high energy capacity: the situation where $e_{\max,1} = 40$ and $e_{\max,2}$ is just a tenth of $e_{\max,1}$ provides a throughput that is 94.3% of that obtained having both sensors with maximum capacity ($e_{\max,i} = 40$). Consequently, if one of the two devices has a relatively large battery, the requirement on the battery of the other device becomes much looser.

In Figure 4.5, the long-term reward versus the harvesting probability of

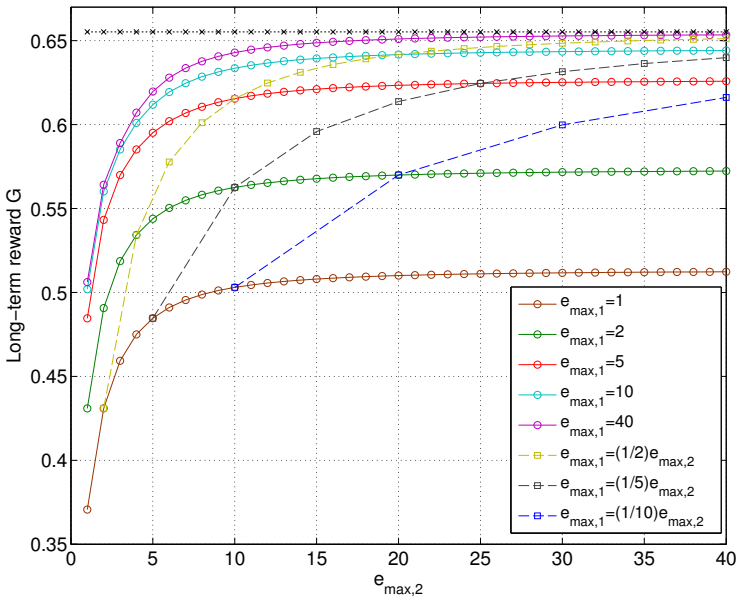

Figure 4.4: Long-term reward versus batteries capacity of EHD 2, for different values of $e_{\max,1}$ and energy harvesting rate $\bar{b} = 0.1$ (upper bound shown as the dotted line)

the second EHD $\bar{b}_2$ is represented, for different values of $e_{\max,1}$. It can be seen that the curves from $e_{\max,1} = 10$ to $e_{\max,1} = 30$ roughly present the same behavior. This behavior is consistent with Fig. 4.1, where the performance quickly saturates as the battery capacity grows. As seen in Figure 4.4, this implies that having just one well-equipped EHD is sufficient to achieve good results for the entire sensor system. Moreover, even if, as expected, the reward increases with the harvesting rate of the second EHD, it must be noticed that, in real scenarios, it is very unlikely to enjoy a constant high rate for a long period of time: consequently the right-most part of the figure is practically infeasible. Nevertheless, the study of practical cycles of harvesting rate is currently ongoing.

Finally, Figure 4.6 shows the long-term reward versus the harvesting rate of the second EHD, $\bar{b}_2$, for different values of $\bar{b}_1/\bar{b}_2$ and $e_{\max,1} = e_{\max,2} = 20$. It can be seen that, when the two EHDs have the same harvesting rate, the reward increases with $\bar{b}_i$ until the value $\bar{b}_1 = \bar{b}_2 = 0.5$ is reached. This is because the two EHDs share a common channel, and only one of them is allowed to transmit at any given time, so that $\bar{b}_1 = \bar{b}_2 = 0.5$ saturates the channel, whereas, when $\bar{b}_1 = \bar{b}_2 > 0.5$, some of the energy is lost because of overflow. This saturation effect cannot be observed in the other cases where $b_1 < b_2$ and $b_1 + b_2 \geq 1$. The reason can be understood by considering the extreme

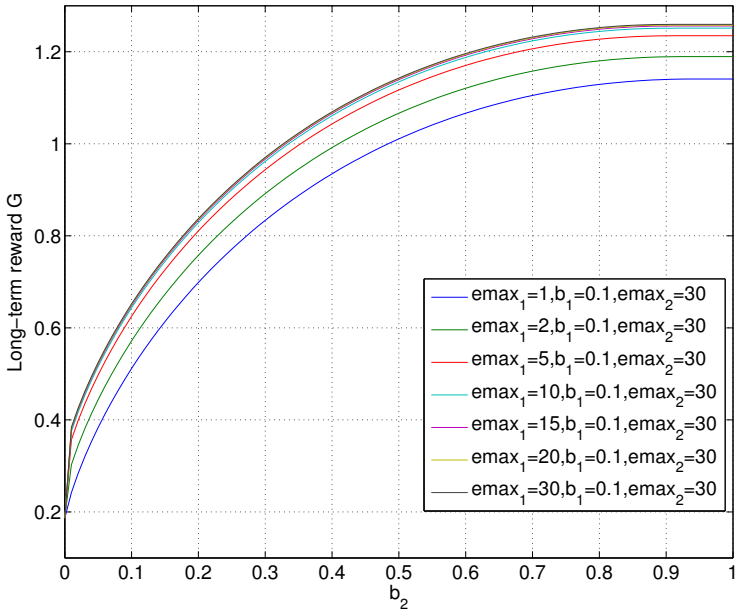

Figure 4.5: Long-term reward versus EH probabilities of the second EHD, for different values of $e_{\max,1}$

scenario $b_1 = 0,\ b_2 = 1$. In this case, only EHD 2 transmits, with probability 1, hence the long-term reward is $\mathbb{E}[V_1]$. On the other hand, when $b_1 = b_2 = 0.5$, assuming the balanced policy $\eta_i(\mathbf{e}) = \bar{b}_i = 0.5$ for both EHDs, the EHD whose packet has maximal importance transmits, so that the asymptotic long-term reward (for $e_{\max,1}, e_{\max,2} \to \infty$) is $\mathbb{E}[\max\{V_1, V_2\}] > \mathbb{E}[V_1]$. It can be concluded that a performance degradation is incurred when the harvesting rates of the two EHDs are unbalanced.

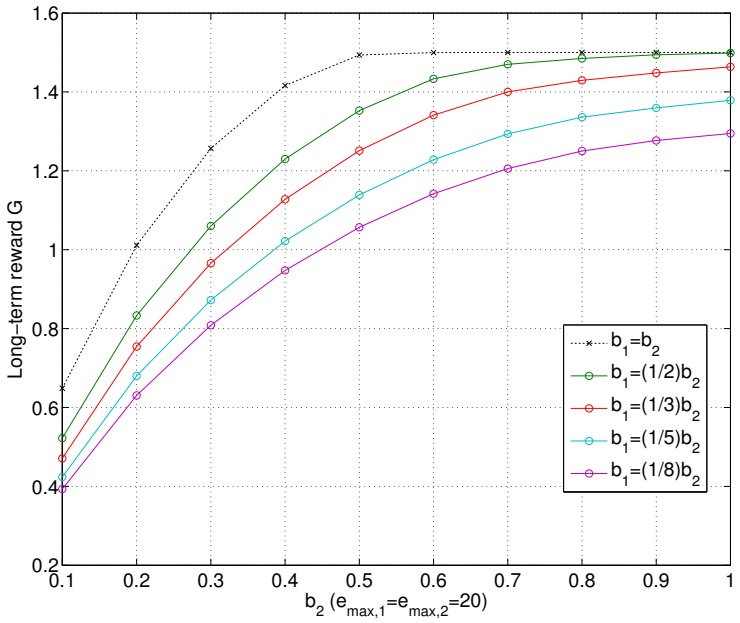

Figure 4.6: G versus EH rate of EHD 2, for $e_{\max,1} = e_{\max,2} = 20$ and b_1/b_2

Chapter 5

Rate maximization in a general scenario

The purpose of this chapter is to resume the framework in Chapter 3, performing the mathematical analysis in a more general setting: the *low SNR regime* scenario will be dropped and the expressions of (3.4), (3.5) and (3.6) will be generalized.
This new environment has the advantage of allowing a more realistic comparison of the performance achieved with different measures of the SNRs of the two EHDs, which will be able to take values in a generic set, but at the cost of an increased complexity both in the mathematical analysis and in the algorithm.

5.1 Variables calculation

Using the exact expression of the achievable rate in slot k for EHD i (3.1):

$$V_{i,k} = \ln(1 + \Lambda_i H_{i,k}),$$

and remembering that a threshold on the importance level $v_{i,\text{th}}$ corresponds to a threshold on the normalized channel gain $h_{i,\text{th}} = \frac{e^{v_{i,\text{th}}}-1}{\Lambda_i}$, exploiting again the independence between the exponentially distributed channels, it can be

obtained:

$$\begin{aligned}\eta_1(\mathbf{e}) &= \mathbb{P}\left[H_{1,k} > h_{1,th}(\mathbf{e}), H_{2,k} < \alpha_1 + \beta_1 H_{1,k}\right] \\ &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\alpha_1+\beta_1 h_1} f_{H_1}(h_1) f_{H_2}(h_2) \mathrm{d}h_2 \mathrm{d}h_1 \\ &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\alpha_1+\beta_1 h_1} e^{-(h_1+h_2)} \mathrm{d}h_2 \mathrm{d}h_1 \\ &= e^{-h_{1,th}(\mathbf{e})} \left(1 - \frac{e^{-\alpha_1-\beta_1 h_{1,th}(\mathbf{e})}}{1+\beta_1}\right), \end{aligned} \tag{5.1}$$

where

$$\begin{aligned}\alpha_1 &= \frac{1}{1+\Lambda_1 h_{1,th}(\mathbf{e})} \left(h_{2,th}(\mathbf{e}) - \frac{\Lambda_1}{\Lambda_2} h_{1,th}(\mathbf{e})\right), \\ \beta_1 &= \frac{1+\Lambda_2 h_{2,th}(\mathbf{e})}{1+\Lambda_1 h_{1,th}(\mathbf{e})} \frac{\Lambda_1}{\Lambda_2};\end{aligned}$$

then, by symmetry,

$$\eta_2(\mathbf{e}) = e^{-h_{2,th}(\mathbf{e})} \left(1 - \frac{e^{-\alpha_2-\beta_2 h_{2,th}(\mathbf{e})}}{1+\beta_2}\right), \tag{5.2}$$

where

$$\begin{aligned}\alpha_2 &= \frac{1}{1+\Lambda_2 h_{2,th}(\mathbf{e})} \left(h_{1,th(\mathbf{e})} - \frac{\Lambda_2}{\Lambda_1} h_{2,th}(\mathbf{e})\right), \\ \beta_2 &= \frac{1+\Lambda_1 h_{1,th}(\mathbf{e})}{1+\Lambda_2 h_{2,th}(\mathbf{e})} \frac{\Lambda_2}{\Lambda_1};\end{aligned}$$

and η_0 maintains its usual expression:

$$\eta_0(\mathbf{e}) = 1 - \eta_1(\mathbf{e}) - \eta_2(\mathbf{e}) = \left(1 - e^{-h_{1,th}(\mathbf{e})}\right)\left(1 - e^{-h_{2,th}(\mathbf{e})}\right).$$

The exact expression of (2.4) is

$$\begin{aligned}&g(\eta_1(\mathbf{e}), \eta_2(\mathbf{e})) = \mathbb{E}[Q_{1,k}V_{1,k} + Q_{2,k}V_{2,k} | (E_{1,k}, E_{2,k}) = \mathbf{e}] \\ &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_0^{\alpha_1(\mathbf{e})+\beta_1(\mathbf{e})h_1} \ln(1+\Lambda_1 h_1) e^{-(h_1+h_2)} \mathrm{d}h_2 \mathrm{d}h_1 \\ &+ \int_{h_{2,th}(\mathbf{e})}^{\infty} \int_0^{\alpha_2(\mathbf{e})+\beta_2(\mathbf{e})h_2} \ln(1+\Lambda_2 h_2) e^{-(h_1+h_2)} \mathrm{d}h_1 \mathrm{d}h_2 \\ &\triangleq g_1 + g_2 \end{aligned} \tag{5.3}$$

$$
\begin{aligned}
g_1 &= \int_{h_{1,th}(\mathbf{e})}^{\infty} \int_{0}^{\alpha_1(\mathbf{e})+\beta_1(\mathbf{e})h_1} \ln(1+\Lambda_1 h_1) e^{-(h_1+h_2)} \mathrm{d}h_2 \mathrm{d}h_1 \\
&= \int_{h_{1,th}(\mathbf{e})}^{\infty} \ln(1+\Lambda_1 h_1) e^{-h_1} dh_1 - e^{-\alpha_1(\mathbf{e})} \left(\int_{h_{1,th}(\mathbf{e})}^{\infty} \ln(1+\Lambda_1 h_1) e^{-h_1(1+\beta_1(\mathbf{e})} dh_1 \right) \\
&= -e^{\frac{1}{\Lambda_1}} Ei\left(-\frac{\Lambda_1 h_{1,th}(\mathbf{e})+1}{\Lambda_1} \right) + e^{-h_{1,th}(\mathbf{e})} \ln(\Lambda_1 h_{1,th}(\mathbf{e})+1) + \\
&+ e^{-\alpha_1(\mathbf{e})} \left[\frac{e^{\frac{1+\beta_1(\mathbf{e})}{\Lambda_1}}}{1+\beta_1(\mathbf{e})} Ei\left(-\frac{(1+\beta_1(\mathbf{e}))(\Lambda_1 h_{1,th}(\mathbf{e})+1)}{\Lambda_1} \right) + \right. \\
&\left. - \frac{e^{-(1+\beta_1(\mathbf{e}))h_{1,th}(\mathbf{e})}}{1+\beta_1(\mathbf{e})} \ln(\Lambda_1 h_{1,th}(\mathbf{e}))+1) \right]
\end{aligned}
$$

and analogously

$$
\begin{aligned}
g_2 &= -e^{\frac{1}{\Lambda_2}} Ei\left(-\frac{\Lambda_2 h_{2,th}(\mathbf{e})+1}{\Lambda_2} \right) + e^{-h_{2,th}(\mathbf{e})} \ln(\Lambda_2 h_{2,th}(\mathbf{e})+1) + \\
&+ e^{-\alpha_2(\mathbf{e})} \left[\frac{e^{\frac{1+\beta_2(\mathbf{e})}{\Lambda_2}}}{1+\beta_2(\mathbf{e})} Ei\left(-\frac{(1+\beta_2(\mathbf{e}))(\Lambda_2 h_{2,th}(\mathbf{e})+1)}{\Lambda_2} \right) + \right. \\
&\left. - \frac{e^{-(1+\beta_2(\mathbf{e}))h_{2,th}(\mathbf{e})}}{1+\beta_2(\mathbf{e})} \ln(\Lambda_2 h_{2,th}(\mathbf{e}))+1) \right]
\end{aligned}
$$

where the calculation has exploited the following

Definition 3. *(Exponential integral)*
For real nonzero values of x, the Exponential Integral Ei(x) is defined as

$$
Ei(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt \tag{5.4}
$$

5.2 Algorithm implementation

Given the complexity of (5.1) and (5.2) compared to (3.4) and (3.5), an analytical expression of $h_{1,th}(\mathbf{e})$ and $h_{2,th}(\mathbf{e})$ as a function of $\eta_i(\mathbf{e})$ is not computable and, as a consequence, the optimization (3.11) over the values of $\eta_i(\mathbf{e})$, as performed in Section 3.3, is not feasible. However, due to the one-to-one mapping between the transmission probabilities and the thresholds of the channels, it is still possible to maximize (2.5) over $h_{i,th}(\mathbf{e})$ and then, by (5.1) and (5.2)

obtain the optimal $\eta_i(\mathbf{e})$.
Therefore, for the unconstrained scenario, the *Policy Improvement* step of PIA has been implemented numerically, using the MATLAB® function `fmincon`, which allows to find the minimum of a constrained nonlinear multivariable function. Among MATLAB®'s optimization Algorithms, namely "Trust Region Reflective", "Active Set", "SQP", "Interior Point", the last has been used, due to the fact that

- it satisfies bounds at all iterations
- it handles large sparse problems, as well as small dense problems
- it produces warnings when NaN or Inf results occur.

Briefly speaking, the interior-point approach to constrained minimization is to solve a sequence of approximate minimization problems: if the original problem is

$$\begin{cases} \min\limits_{x} f(x) \\ h(x) = 0 \\ g(x) \leq 0 \end{cases} \tag{5.5}$$

then, $\forall \mu > 0$, the approximate problem is

$$\begin{cases} \min\limits_{x,s} f_\mu(x,s) = \min\limits_{x,s} f(x) - \mu \sum\limits_{i} \ln(s_i) \\ h(x) = 0 \\ g(x) + s = 0 \end{cases} \tag{5.6}$$

It can be seen that as many slack variables s_i as inequality constraints g are added, with s_i restricted to be positive to keep $\ln(s_i)$ (called *barrier function)* bounded. As μ decreases to zero, the minimum of f_μ should approach that of f. Consequently, as the approximate problem (5.6) is a sequence of equality constrained problems rather than an inequality-constrained problem as (5.5), it is easier to solve [40].

Chapter 6

Numerical results in the unconstrained scenario

Figure 6.1 shows the long-term reward for $e_{max,1} = e_{max,2} = 20$ and $\bar{b}_1 = \bar{b}_2 = 0.1$ in the case where the reward g is calculated by (3.6) or (5.3). Since the relation between the two expressions is that the first is the linear approximation of the second, it can be seen that their behavior is the same only for small values of $\Lambda_1 = \Lambda_2 = \Lambda$. In fact, although for $\Lambda = 0.1$ the approximated value is just 16% greater than the exact one, for $\Lambda = 1$ the ratio is 129% which becomes even 478% for $\Lambda = 5$.

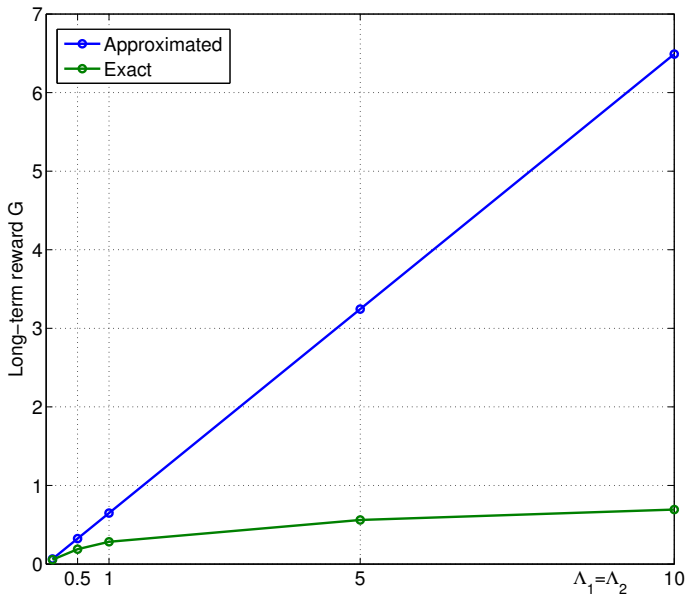

Figure 6.1: Comparison between the values of G calculated using (3.6) and (5.3)

In Figure 6.2 the long-term reward G for different values of the Λ_1 is plotted, as a function of the batteries capacity $e_{max,1} = e_{max,2}$. It can be noted that the throughput saturates for $e_{max,1} = e_{max,1} \simeq 10$, which is analogous to what noted for the *low SNR* scenario of Chapter 3. Moreover, the result obtained when just an EHD benefits from a high SNR is much better than that achieved with an equal-SNR scenario: the reward for $\Lambda_1 = 20\Lambda_2$ is twice that for $\Lambda_1 = \Lambda_2$. Consequently this means that, if at least one EHD is in a good transmission environment, the obtained reward can be high, even if probably come from a single node.

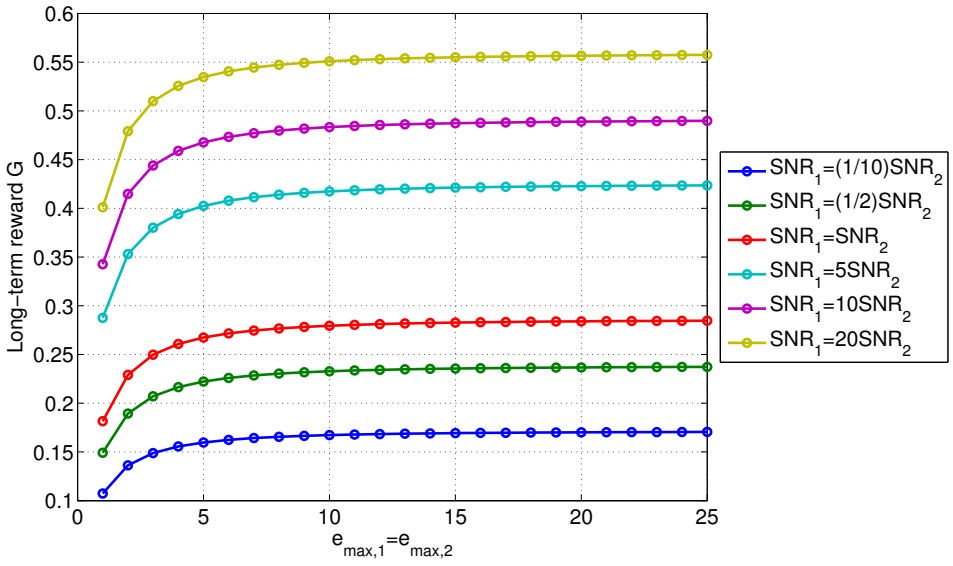

Figure 6.2: Long-term reward G for different values of the Λ_1 and $\Lambda_2 = 1$

The optimal policy is depicted in Figure 6.3, for $e_{max,1} = e_{max,2} = 40$ and $b_1 = b_2 = 0.1$. It can be seen that it is conservative in the low energy levels (the value of η_1 and η_2 is low in the first states), thus avoiding *energy outage*, and aggressive in the higher ones (η_1 and η_2 increase as energy becomes abundant), thus avoiding *energy overflow*. Moreover, another important fact is that there is a "symmetry" between the values assumed by the two transmission probabilities: $\eta_1(x, y) = \eta_2(y, x)$, $\forall(x, y)$ (this occurs only when $e_{max,1} = e_{max,2}$, $\Lambda_1 = \Lambda_2$ and $b_1 = b_2 = 0.1$, *i.e.*, the system is symmetric). Finally, it results that η_1 mainly depends on e_1 and, for a fixed value of e_1, is almost independent of e_2 (and analogously for η_2): this observation can be exploited to create some approximated lower-complexity solutions, like the one presented in the next Section.

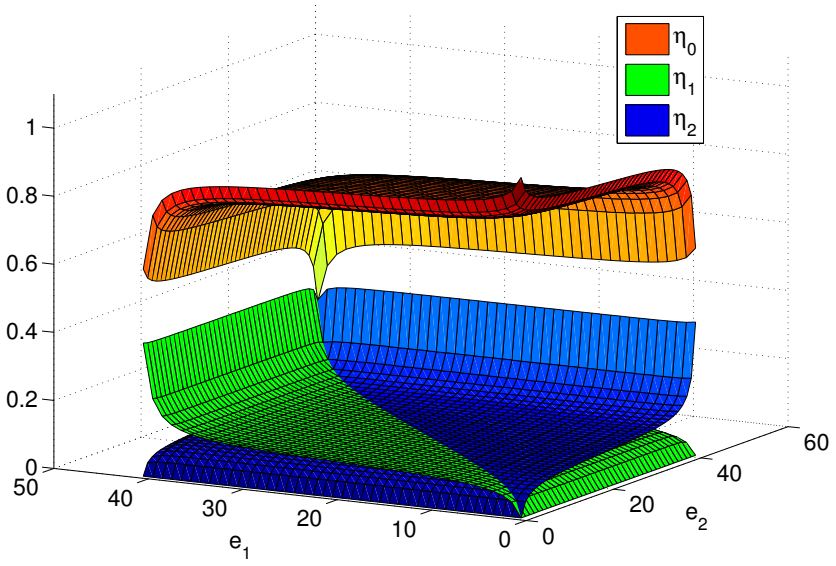

Figure 6.3: η_0,η_1 and η_2 for the optimal policy ($e_{max,1} = e_{max,2} = 40$)

6.1 Low-complexity policy

Usually the PIA is computationally intensive for the low-power electronics typically present in practical EHDs. As a consequence, the design of low-complexity policies, able to achieve almost optimal performance, can be an important step towards an energy-efficient realistic implementation of an optimal transmission strategy. If the framework parameters are expected to be mainly constant, it is possible to directly store the transmission probabilities $\eta_i(\mathbf{e})$, previously determined offline, in a register of the devices, so as to completely avoid the energy-draining PIA. However, if the policy is complex and the size of the batteries is large, the number of elements to be stored becomes too big to be contained in such register. As a result, a possible solution is that of utilizing approximated policies which, although using a smaller number of elements, still achieve good performance.
In this section an example of this approach is described: exploiting the last two remarks about the optimal policy of Figure 6.3, it is possible to develop a transmission strategy that:

- for every value of e_1, substitutes $\eta_1(e_1, e_2)$, $e_2 = 0, \ldots, e_{max,2}$ with the mean value of $\eta_1(e_1, 0), \eta_1(e_1, 1), \eta_1(e_1, 2), \ldots, \eta_1(e_1, e_{max,2})$
- for every value of e_2, substitutes $\eta_2(e_1, e_2)$, $e_1 = 0, \ldots, e_{max,2}$ with the mean value of $\eta_2(0, e_2), \eta_2(1, e_2), \eta_2(2, e_2), \ldots, \eta_2(e_{max,1}, e_2)$

- sets the values of $\eta_0(e_1, e_2)$ as $1 - \eta_1(e_1, e_2) - \eta_2(e_1, e_2),\ \forall e_1, e_2$.

The results obtained with this approach are depicted in Figure 6.4, which shows the difference between the long-term reward G achieved by the optimal policy and that obtained by the LCP, as a function of $e_{max,1} = e_{max,2} = e_{max}$.

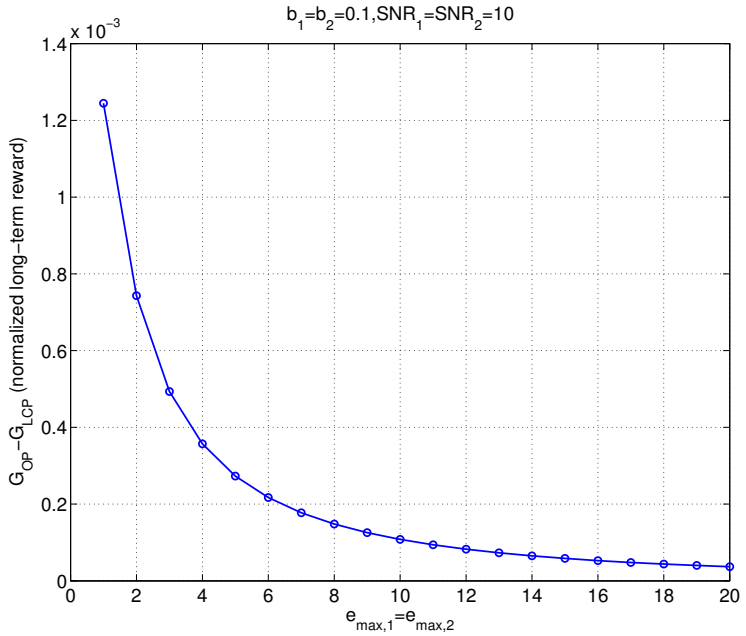

Figure 6.4: $G_{OP} - G_{LCP}$ for $b_1 = b_2 = 0.1$ and $SNR_1 = SNR_2 = 10$

It can be seen that the LCP already performs very well even for low values of e_{max}, and almost overlaps with the optimal policy as this parameter increases. The memory savings can be estimated considering that, while the optimal policy saves $3(e_{max} + 1)^2$ different values, the LCP only stores $\frac{(e_{max}+1)(e_{max}+4)}{2}$ elements, with a reduction in the need for memory space of roughly 80%.

Finally, the significant independence of the energy level of a node from that of the other suggests the possibility to develop heuristic distributed transmission strategies, to be studied in a future research activity.

Chapter 7

Dealing with Imperfect State-of-Charge Knowledge

So far, the implicit assumption of perfect knowledge of the energy available in the EHDs has been made. However, estimating the energy level of the batteries or super-capacitors employed in real-world EHDs, commonly known as State-Of-Charge (SOC), is actually a non-trivial task. This Chapter deals with the design of an EHD operation policy in which the central controller has only imperfect knowledge of the available amount of energy stored in the two sensors. In [41], it is stated that variations in the super-capacitor capacitance relative to the data-sheet value, due to age or temperature fluctuations, may be of the order of 30%. An online SOC estimation algorithm based on controlled discharge is thus proposed and shown to perform well, albeit at a small energy loss. In [42] and [43] different high-complexity algorithms are designed for the estimation of the open circuit voltage of an electrochemical battery, which is linearly related to the SOC, showing that precise knowledge of the SOC may be unreliable or too expensive. Motivated by the aforementioned real-world concerns, this Chapter extends [33], which is devoted to the analysis and design of optimal energy management policies for the imperfect-SOC scenario of a single EHD, carrying out the mathematical analysis and identifying new results, with the aim of characterizing the more general scenario consisting in two EHDs.

7.1 System model

The system model for the imperfect-SOC scenario is similar to that described in Chapter 2. The evolution at time k of the amount of energy quanta $E_{i,k}$ at

EHD i is now determined by

$$E_{i,k+1} = \min\{[E_{i,k} - Q_{i,k}]^+ + B_{i,k}, e_{max,i}\} \tag{7.1}$$

where:

- $\{B_{i,k}\}$, the *energy arrival process*, takes values in $\mathcal{B}_i = \{0, 1, \ldots, b_{max,i}\}$ and has a geometric probability mass function $p_{B_i}(b),\ b \in \mathcal{B}_i$, with mean $\bar{b_i}$;

- $Q_{i,k}$ is the number of energy quanta requested by the central controller to EHD i, in slot k. $\mathcal{Q}_i = \{0, \ldots, q_{max,i}\}$ is the *action space* related to EHD i, for some $0 < q_{max,i} \leq e_{max,i}$, so that $Q_{i,k} \in \mathcal{Q}_i, \forall k$. Finally, the joint energy requested from CC to the couple (EHD1,EHD2) at time k is $\mathbf{Q}_k$, with $\mathbf{Q}_k \in \mathcal{Q}$:

  $$\mathcal{Q} = \{(0,0), (0,1), \ldots, (0, q_{max,2}), (1,0), \ldots, (q_{max,1}, 0)\}, \tag{7.2}$$

  $\forall k$. The parameter $q_{max,i}$ reflects a physical constraint on the maximum amount of energy that can be drawn from the buffer at any given time. Clearly, due to the usual collision model employed and the centralized controller, $\mathcal{Q}$ cannot include (q_1, q_2) with q_1 and q_2 simultaneously positive.

Assuming that the two EHDs are symmetric ($e_{max,1} = e_{max,2} = e_{max}$, $\mathcal{B}_1 = \mathcal{B}_2,\ p_{B_1}(b) = p_{B_2}(b)$ and $\mathcal{Q}_1 = \mathcal{Q}_2$) and only partial knowledge of $E_{i,k}$ is available at the controller, let $\mathcal{I}(n),\ n = \{0, 1, 2, 3\}$ be a partition of the state space $\mathcal{E}^2$, defined as

- $\mathcal{I}(0) = \{(e_1, e_2),\ 0 \leq e_1, e_2 \leq \tilde{e} - 1\}$ ("LL")
- $\mathcal{I}(1) = \{(e_1, e_2),\ 0 \leq e_1 \leq \tilde{e} - 1,\ \tilde{e} \leq e_2 \leq e_{max}\}$ ("LH")
- $\mathcal{I}(2) = \{(e_1, e_2),\ \tilde{e} \leq e_1 \leq e_{max},\ 0 \leq e_2 \leq \tilde{e} - 1\}$ ("HL")
- $\mathcal{I}(3) = \{(e_1, e_2),\ \tilde{e} \leq e_1, e_2 \leq e_{max}\}$ ("HH")

with $\tilde{e} = \lceil \frac{e_{max}}{2} \rceil$. The assumption is that, at time k, $(E_{1,k}, E_{2,k}) \in \mathcal{I}(N_k)$, with $N_k \in \{0, 1, 2, 3\}$, and the central controller knows only that $\mathbf{E}_k = (E_{1,k}, E_{2,k}) \in \mathcal{I}(N_k)$, *i.e.*, N_k rather than the exact SOC $\mathbf{E}_k$.

The definitions of energy outage and overflow have to be updated in the following way:

Definition 4. *Energy outage occurs, for EHD i, when $Q_{i,k} > E_{i,k}$, as a consequence of the imperfect knowledge of $E_{i,k}$, due to which the controller may attempt to draw more energy than what is available*

Definition 5. *Energy overflow occurs if $B_{i,k} > e_{max} - [E_{i,k} - Q_{i,k}]^+$,* i.e., *the energy buffer is unable to store all of the harvested energy $B_{i,k}$.*

7.2 Policy definition and optimization problem

Given the interval index N_k, a policy μ decides on the amount of energy $\mathbf{Q}_k$ to be drawn from the pair of sensors. Formally, μ is a probability measure on the action space $\mathcal{Q}$, parameterized by N_k, *i.e.*, given N_k, $\mu(\mathbf{q}; N_k)$ is the probability of choosing action $\mathbf{q} \in \mathcal{Q}$ in slot k.

The reward $g : \mathcal{Q} \times \mathcal{E}^2 \to \mathbb{R}^+$ is defined as

$$g(Q_{1,k}, Q_{2,k}, E_{1,k}, E_{2,k}) = \begin{cases} 0 & Q_{1,k} > E_{1,k} \text{ or } Q_{2,k} > E_{2,k} \\ \tilde{g}(Q_{1,k}, Q_{2,k}) & Q_{1,k} \leq E_{1,k} \text{ and } Q_{2,k} \leq E_{2,k} \end{cases} \tag{7.3}$$

where $\tilde{g} : \mathcal{Q} \to \mathbb{R}^+$ is a concave increasing function of $Q_{1,k}$ and $Q_{2,k}$, with $\tilde{g}(0,0) = 0$.
When $Q_{i,k} > E_{i,k}$ the reward is 0, which models the inability of the sensor node i to complete the requested task, when there is energy outage. As an example, if the reward function is the transmission rate, then, according to the Shannon formula, $\tilde{g}(Q_{1,k}, Q_{2,k}) = \ln(1 + \alpha \bar{Q}_k)$, where $\alpha > 0$ is an SNR scaling factor and $\bar{Q}_k$ is the positive element of the couple $(Q_{1,k}, Q_{2,k})$. The controller spreads the energy $\bar{Q}_k$ over the entire codeword. If $\bar{Q}_k$ is greater than the corresponding value of E_k, that EHD runs out of energy when only a fraction of the codeword has been transmitted, hence the codeword is discarded.
Given $\mathbf{E}_0 = (E_{1,0}, E_{2,0})$, the *long-term average reward* per time-slot under policy μ is defined as

$$G(\mu, \mathbf{E}_0) = \lim_{K \to \infty} \inf \frac{1}{K} \mathbb{E}\left[\sum_{k=0}^{K-1} g(Q_{1,k}, Q_{2,k}, E_{1,k}, E_{2,k}) \middle| \mathbf{E}_0 \right] \tag{7.4}$$

where the expectation is over $\{B_{i,k}, Q_{i,k},\ k = 0, \dots, K-1\}$.
The general problem is to obtain a policy μ^* such that

$$\mu^* = \arg\max_{\mu} G(\mu, \mathbf{E}_0) \tag{7.5}$$

Under SOC uncertainty, by definition of the policy μ, $\mathbf{Q}_k$ is the same for all $(E_{1,k}, E_{2,k}) \in \mathcal{I}(N_k)$. This constraint is not linear, hence (7.5) cannot be solved via standard optimization techniques: in this work an exhaustive search method is thus employed. Furthermore, in order to reduce the complexity, only deterministic policies are considered, in which

$$\begin{cases} \mu_\rho(\mathbf{q}; n) = 1 & \mathbf{q} = \rho(n) \\ \mu_\rho(\mathbf{q}; n) = 0 & \mathbf{q} \in \mathcal{Q} \setminus \{\rho(n)\} \end{cases} \tag{7.6}$$

where $\rho : \{0, 1, 2, 3\} \to \mathcal{Q}$ is a function mapping the interval index $n \in \{0, 1, 2, 3\}$ to the action $\mathbf{q} \in \mathcal{Q}$. As a result, (7.5) can equivalently be written as

$$\rho^* = \arg\max_{\rho} G(\mu, \mathbf{E}_0), \tag{7.7}$$

where, from (7.4),

$$G(\mu_\rho, \mathbf{E}_0) = \sum_{n=0}^{3} \sum_{\mathbf{e} \in \mathcal{I}(n)} \pi_\rho(\mathbf{e}; \mathbf{e}_0) g(\rho(n), \mathbf{e}) \tag{7.8}$$

and $\pi_\rho(\mathbf{e}; \mathbf{e}_0)$ is the asymptotic distribution of the SOC $\mathbf{e} \in \mathcal{E}^2$, given that the initial state is $\mathbf{E}_0 = \mathbf{e}_0$:

$$\pi_\rho(\mathbf{e}; \mathbf{e}_0) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} P_\rho(\mathbf{E}_k = \mathbf{e} | \mathbf{E}_0 = \mathbf{e}_0), \tag{7.9}$$

where $P_\rho(\mathbf{E}_k = \mathbf{e} | \mathbf{E}_0 = \mathbf{e}_0)$ is the k-step transition probability of the chain under policy ρ. The asymptotic distribution can be evaluated as the unique solution of the system of steady-state equations

$$\begin{cases} \displaystyle\sum_{n=0}^{3} \sum_{\mathbf{e} \in \mathcal{I}(n)} \pi_\rho(\mathbf{e}) = 1 & \text{(normalization)} \\ \pi_\rho(\mathbf{e}) \geq 0,\ \forall \mathbf{e} \in \mathcal{E}^2 & \text{(non-negativity)} \\ \displaystyle\sum_{n=0}^{3} \sum_{\mathbf{s} \in \mathcal{I}(n)} \pi_\rho(\mathbf{s}) P_\rho(\mathbf{E}_1 = \mathbf{e} | \mathbf{E}_0 = \mathbf{s}) = \pi_\rho(\mathbf{e}),\ \forall \mathbf{e} \in \mathcal{E}^2 & \text{(steady-state equations)} \end{cases}$$

and so π_ρ is independent of the initial state $\mathbf{E}_0 = \mathbf{e}_0$.

In the next Section, the optimal long-term reward is numerically determined for the particular case considered so far, comparing it with that obtained by a controller with perfect SOC knowledge.

7.3 Numerical results

This section deals with the maximization of (7.5) for $\bar{b} = 20$ and a geometric energy arrival distribution truncated at $b_{max} = 4\bar{b}$. The reward function is the normalized throughput

$$\tilde{g}(Q_{1,k}, Q_{2,k}) = \frac{\ln(1 + \alpha \bar{Q}_k)}{\ln(1 + \alpha \bar{b})}, \tag{7.10}$$

where α is an SNR scaling factor and $\bar{Q}_k$ is the positive element of the couple $(Q_{1,k}, Q_{2,k})$. As a result of the partition performed, the state space is subdivided into 4 regions, namely LL,LH,HL,HH.
As regards the throughput obtained by the optimization performed with perfect SOC knowledge, in this case the controller selects action $\mathbf{Q}_k = \mathbf{q}$ when the SOC is $\mathbf{E}_k$, $\forall \mathbf{E}_k \in \mathcal{E}^2$. Note that the long-term reward under perfect SOC knowledge represents an upper bound to the performance for the imperfect SOC case. Figure 7.1 shows the throughput G versus the battery capacities

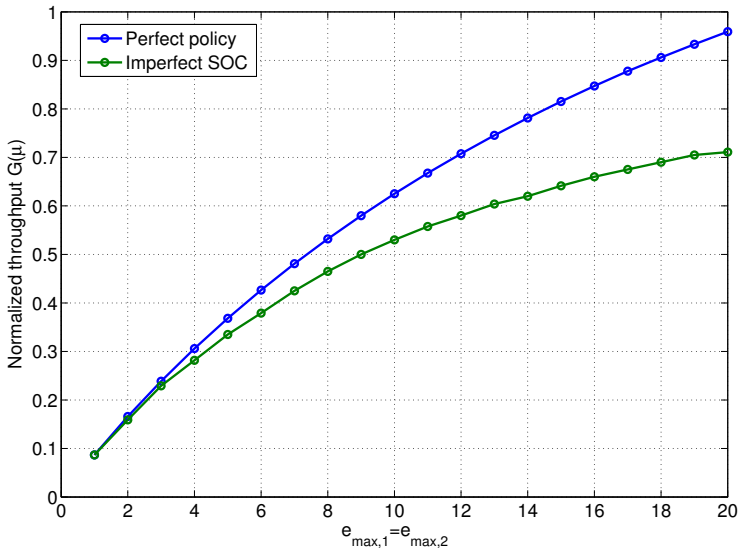

Figure 7.1: Normalized throughput as a function of $e_{max,i}$, with $\alpha = 0.1$, $\bar{b} = 20$

of the devices: it can be seen that, as e_{max} increases, the degradation of the imperfect policy with respect to the perfect policy becomes bigger. This is due to several factors, like the growing number of states included in each of the four partitions of the state space, the choice of not taking into account the previous history of the EHDs pair, and the utilization of a deterministic policy between the energy states and the energy quanta requested by the controller.

The elimination of some of these assumptions will be the goal of future research activities.

Chapter 8

Conclusions

This thesis has studied the case of a Wireless Sensor Network (WSN) consisting of two EHDs which report data of different "importance" levels to a receiver (RX), with the overall goal to maximize the long-term aggregate average importance of the reported data.

In the case of perfect knowledge of the State-Of-Charge (SOC) of EHDs batteries, it has been shown that, for the class of binary (transmit/no transmit) policies, the optimal policy dictates the transmission of data with importance above a given threshold, which is a function of the joint energy levels available in the two devices. Numerical results have been provided, to evaluate the impact on the performance of factors such as the battery capacity size and the energy harvesting rate.

Finally, motivated by real-world EHD implementations, it has been investigated the performance of a transmission policy for a central controller operating under imperfect SOC knowledge. Simulation results for a simple four-state controller, only knowing if the SOC of each node is LOW or HIGH, have shown that the obtained reward may be significantly lower than that achieved by a policy with perfect SOC knowledge.
Future work will further investigate how the energy buffer capacities and the energy arrival distribution influence the performance of EHDs operation policies with imperfect SOC knowledge.

Bibliography

[1] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*. ACM, 2002, pp. 88–97. [Online]. Available: http://dl.acm.org/citation.cfm?id=570751

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128601003024

[3] C. Perkins, *Ad hoc networks*, M. Reading, Ed. Addison-Wesley, 2000.

[4] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*. ACM, 1999, pp. 263–270. [Online]. Available: http://dl.acm.org/citation.cfm?id=313556

[5] J. K. Hart and K. Martinez, "Environmental sensor networks: A revolution in the earth system science?" *Earth-Science Reviews*, vol. 78, no. 3, pp. 177–191, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0012825206000511

[6] T. Le Dinh, W. Hu, P. Sikka, P. Corke, L. Overs, and S. Brosnan, "Design and deployment of a remote robust sensor network: Experiences from an outdoor water quality monitoring network," in *32nd IEEE Conference on Local Computer Networks, 2007. LCN 2007.* IEEE, 2007, pp. 799–806. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4367918

[7] N. Bulusu, D. Estrin, L. Girod, and J. Heidemann, "Scalable coordination for wireless sensor networks: self-configuring localization systems,"

in *International Symposium on Communication Theory and Applications (ISCTA 2001), Ambleside, UK*, 2001.

[8] E. Shih, S.-H. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan, "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proceedings of the 7th annual international conference on Mobile computing and networking*. ACM, 2001, pp. 272–287. [Online]. Available: http://dl.acm.org/citation.cfm?id=381703

[9] E. M. Petriu, N. D. Georganas, D. C. Petriu, D. Makrakis, and V. Z. Groza, "Sensor-based information appliances," *Instrumentation & Measurement Magazine, IEEE*, vol. 3, no. 4, pp. 31–35, 2000. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=887458

[10] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao, "Habitat monitoring: Application driver for wireless communications technology," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2 supplement, pp. 20–41, 2001. [Online]. Available: http://dl.acm.org/citation.cfm?id=844196

[11] G. Hoblos, M. Staroswiecki, and A. Aitouche, "Optimal design of fault tolerant sensor networks," in *Proceedings of the 2000 IEEE International Conference on Control Applications, 2000.* IEEE, 2000, pp. 467–472. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=897468

[12] J. Rabaey, J. Ammer, J. da Silva Jr, and D. Patel, "Picoradio: Ad-hoc wireless networking of ubiquitous low-energy sensor/monitor nodes," in *Proceedings of the 2000 IEEE Computer Society Workshop on VLSI.* IEEE, 2000, pp. 9–12. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=844522

[13] J. M. Rabaey, M. J. Ammer, J. L. da Silva Jr, D. Patel, and S. Roundy, "Picoradio supports ad hoc ultra-low power wireless networking," *Computer*, vol. 33, no. 7, pp. 42–48, 2000. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=869369

[14] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000. [Online]. Available: http://dl.acm.org/citation.cfm?id=332838

[15] E. H. Journal, *Glossary: Energy harvesting.* [Online]. Available: http://www.energyharvestingjournal.com/glossary/energy-harvesting-332.asp

[16] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications.*, vol. 4, no. 5, pp. 2349–2360, 2005. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1532220

[17] A. Sinha and A. Chandrakasan, "Dynamic power management in wireless sensor networks," *Design & Test of Computers, IEEE*, vol. 18, no. 2, pp. 62–74, 2001. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=914626

[18] S. J. Baek, G. De Veciana, and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *IEEE Journal on Selected Areas in Communications.*, vol. 22, no. 6, pp. 1130–1140, 2004. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1321225

[19] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 51–60, 2002. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=985684

[20] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking.* ACM, 1998, pp. 181–190. [Online]. Available: http://dl.acm.org/citation.cfm?id=288286

[21] S. Ratnaraj, S. Jagannathan, and V. Rao, "OEDSR: Optimized energy-delay sub-network routing in wireless sensor network," in *Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, 2006. ICNSC'06.* IEEE, 2006, pp. 330–335. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1673167

[22] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and*

Communications. IEEE Societies, vol. 3. IEEE, 2003, pp. 1713–1723. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1209194

[23] P. Nuggehalli, V. Srinivasan, and R. R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2002, pp. 1765–1772. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1019430

[24] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5441354

[25] Z. Wan, Y. Tan, and C. Yuen, "Review on energy harvesting and energy management for sustainable wireless sensor networks," in *2011 IEEE 13th International Conference on Communication Technology (ICCT)*. IEEE, 2011, pp. 362–367. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6157897

[26] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *Pervasive Computing, IEEE*, vol. 4, no. 1, pp. 18–27, 2005. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1401839

[27] M. Tentzeris. [Online]. Available: http://www.news.gatech.edu/2011/07/06/ambient-electromagnetic-energy-harnessed-small-electronic-devices

[28] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135–1149, 2002. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=995554

[29] R. Arroyo-Valles, A. G. Marques, and J. Cid-Sueiro, "Optimal selective forwarding for energy saving in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 164–175, 2011. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5618890

[30] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Information Theory and Applications Workshop (ITA), 2012*. IEEE, 2012, pp. 249–254. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6181793

[31] N. Michelusi and M. Zorzi, "Optimal random multiaccess in energy harvesting wireless sensor networks," in *Proc. IEEE International Conference on Communications*, 2013.

[32] D. Del Testa, N. Michelusi, and M. Zorzi, "On optimal transmission policies for energy harvesting devices: the case of two users," *IEEE ISWCS 2013*, 2013. [Online]. Available: http://www.vde-verlag.de/proceedings-en/453529076.html

[33] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for energy harvesting devices with imperfect state-of-charge knowledge," in *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 5782–5787. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6364958

[34] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Wireless Networks*, vol. 15, no. 5, pp. 619–635, 2009. [Online]. Available: http://link.springer.com/article/10.1007/s11276-007-0091-0

[35] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2116–2126, 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5504612

[36] K. W. Ross, "Randomized and past-dependent policies for Markov decision processes with multiple constraints," *Operations Research*, vol. 37, no. 3, pp. 474–477, 1989. [Online]. Available: http://or.journal.informs.org/content/37/3/474.short

[37] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2988–3001, 2013. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6522422

[38] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, 1995, vol. 1, no. 2.

[39] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 547–551, 2009. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4786405

[40] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[41] C. Renner and V. Turau, "Caplibrate: self-calibration of an energy harvesting power supply with supercapacitors," in *2010 23rd International Conference on Architecture of Computing Systems (ARCS)*. VDE, 2010, pp. 1–10. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5759026

[42] J. Chiasson and B. Vairamohan, "Estimating the state of charge of a battery," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 3, pp. 465–470, 2005. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1424024

[43] B. Xiao, Y. Shi, and L. He, "A universal state-of-charge algorithm for batteries," in *Proceedings of the 47th Design Automation Conference*. ACM, 2010, pp. 687–692. [Online]. Available: http://dl.acm.org/citation.cfm?id=1837449

Ringraziamenti

Al termine di questo lavoro e del mio ciclo di studi universitari desidero ringraziare i miei genitori, a cui è dedicata questa tesi, per il costante sostegno emotivo ed economico.
Grazie al professor Michele Zorzi e a Nicolò per aver chiarito i miei dubbi e per essere sempre stati disponibili malgrado i numerosi impegni lavorativi.
Un enorme grazie va inoltre ai miei compagni di studi, con cui ho condiviso le mie giornate a "casa DEI", alternando i momenti di studio frenetico a quelli di svago. In particolare, mi sento in dovere di nominare Marco e Giulio, che sempre mi hanno spinto a non arrendermi, aiutandomi a superare le mie difficoltà malgrado le distanze continentali o transoceaniche. Grazie per avermi sempre incoraggiato a dare il massimo: una buona parte dei miei risultati la devo a voi.
Grazie a chi mi è stato vicino al di fuori dell'ambito universitario, e che non mi ha mai fatto mancare il proprio affetto o la propria simpatia: parenti, compagni dell'USMI (con menzione speciale per Lucilla e Luigi) e amici vari della Parrocchia.
Ultima ma più importante di tutte, devo ringraziare quella persona speciale che da ormai tre anni mi sopporta: grazie Elena per essere sempre presente per me, consolandomi nei momenti di sconforto e rallegrandoti con me in quelli di gioia.
Infine, grazie al mio nemico di sempre MATLAB: malgrado quattro anni e mezzo di reciproca antipatia siamo finalmente riusciti ad andare (quasi) d'accordo.